Robust Power System State Estimation for the Nonlinear AC Flow Model

Hao Zhu, Student Member, IEEE, and Georgios B. Giannakis, Fellow, IEEE

Abstract—An important monitoring task for power systems is accurate estimation of the system operation state. Under the nonlinear AC power flow model, the state estimation (SE) problem is inherently nonconvex giving rise to many local optima. In addition to nonconvexity, SE is challenged by data integrity and cyber-security issues. Unfortunately, existing robust (R-) SE schemes employed routinely in practice rely on iterative solvers, which are sensitive to initialization and cannot ensure global optimality. A novel R-SE approach is formulated here by capitalizing on the sparsity of an overcomplete outlier vector model. Observability and identifiability issues of this model are investigated, and neat links are established between R-SE and error control coding. The convex semidefinite relaxation (SDR) technique is further pursued to render the nonconvex R-SE problem efficiently solvable. The resultant algorithm markedly outperforms existing iterative alternatives, as corroborated through numerical tests on the standard IEEE 30-bus system.

Index Terms—Power system state estimation, robustness, sparsity, system identifiability, semidefinite relaxation.

I. INTRODUCTION

The electric power grid is a complex cyber-physical system consisting of multiple modules, each with a transmission infrastructure spanning over a huge geographical area, transporting energy from generation sites to distribution networks. Monitoring the operational conditions of grid transmission networks is of paramount importance to facilitate system control and optimization tasks, including security analysis and economic dispatch with security constraints; see e.g., [1] Ch. 1] and [16]. For this purpose, various system variables are measured in distant buses and then transmitted to the control center for estimating the state variables, namely complex bus voltages. Due to the wide spread of transmission networks and the current integration of enhanced computer/communication infrastructure, the power system state estimation (SE) is challenged by data integrity concerns arising due to “anomalous” measurements affected by outliers [12], [20] and/or adversarial cyber-attacks [4], [13], [14]. These concerns motivate the development of robust approaches to improve resilience of SE to anomalous (a.k.a. bad) data.

For the AC power flow model however, SE challenges come not only from anomalous data, but are further magnified due to the nonlinear couplings present between meter measurements and state variables. To cope with these challenges, Gauss-Newton iterative solvers estimate the state of an approximate linear regression model per iteration, using robust renditions of the weighted least-squares (WLS) error criterion, such as the weighted least-absolute value (WLAV) one; see e.g., [1] Ch. 2-6]. The current iteration adopts Taylor’s first-order expansion around the estimate of the previous iteration to approximate the quadratic AC model with the aforementioned linear regression model. This iterative procedure is closely related to gradient descent algorithms for solving nonconvex problems, see e.g., [3] Ch. 1], which are known to encounter two issues: i) sensitivity to initialization; and ii) convergence concerns. Existing variants have asserted improved numerical stability of the iterative procedure [1] Ch. 3]. Latest SE trends incorporate linear state measurements offered by synchronized phasor measurement units (PMUs); see e.g., [10] and references therein. However, limited PMU deployment currently confines SE to mostly rely on the traditional nonlinear meter measurements, and its companion Gauss-Newton iterative methods. Hence, it is very important to develop a robust (R-) SE solver tailored for the nonlinear measurement model, and capable of approximating the global optimum at polynomial complexity.

The present paper adopts an overcomplete additive outlier-aware measurement model, and leverages the sparsity of outliers to develop an R-SE approach to jointly estimate system states and identify the outliers present (Section II).

Inherent to the overcomplete outlier-aware model is system under-determinacy, which in turn raises outlier observability and identifiability concerns (Section III). It was recently recognized that there are unobservable cyber-attacks that the system operator would fail to detect [3], [13], [14], but all studies so far are limited to linear approximate SE models. Theoretical guarantees of the sparse outlier model were explored for nonlinear SE models in [20]. Compared to these works, the present one provides a unifying framework to understand how tolerant the nonlinear regression model is to data corruption, by introducing the notion of measurement distance. The latter is nicely connected to distance metrics popular in channel coding theory, which are known to determine the error-control capability of channel codes. This connection reveals why the measurement distance is instrumental to characterizing the regression function’s resilience to outliers.

In addition, the novel R-SE framework lends itself to a convex relaxation approach, which yields R-SE solvers approximating the global optimum (Section IV). A well appreciated tool for convexifying non-convex problems [15], semidefinite relaxation (SDR) solvers thus emerge as powerful schemes for R-SE of nonlinear AC power flow models. Preliminary tests on the IEEE 30-bus system corroborate the performance...
improvement of the proposed approach (Section V).

Notation: Upper (lower) boldface letters will be used for matrices (column vectors); \((\cdot)^T\) denotes transposition; \((\cdot)^H\) complex-conjugate transposition; \(\text{Re}(\cdot)\) the real part; \(\text{Im}(\cdot)\) the imaginary part; \(\text{Tr}(\cdot)\) the matrix trace; \(\text{rank}(\cdot)\) the matrix rank; \(\mathbf{0}\) the all-zero matrix; \(|\cdot|_p\) the vector \(p\)-norm for \(p \geq 1\); \(|\cdot|\) the floor of a real number; and \(|\cdot|\, (\angle)\) the magnitude (angle) of a complex number.

II. MODELING AND PROBLEM STATEMENT

Consider a power transmission network with \(N\) buses denoted by the set of nodes \(\mathcal{N} := \{1, \ldots, N\}\), and \(L\) transmission lines represented by the set of edges \(\mathcal{E} := \{(n, m)\} \subseteq \mathcal{N} \times \mathcal{N}\). Suppose \(M\) measurements are taken for estimating the complex voltage states \(\{V_n\}_{n \in \mathcal{N}}\), from a subset of the following system variables:

- \(P_n(Q_n)\): the real (reactive) power injection at bus \(n\) (negative if bus \(n\) is connected to a load);
- \(P_{mn}(Q_{mn})\): the real (reactive) power flow from bus \(m\) to bus \(n\); and
- \(|V_n|\): the voltage magnitude at bus \(n\).

Compliant with the AC power flow model [2], these measurements obey nonlinear equations relating them with the system state vector \(\mathbf{v} := [V_1, \ldots, V_N]^T \in \mathbb{C}^N\). These equations also involve the injected currents of all buses that are here collected in the vector \(\mathbf{i} := [I_1, \ldots, I_N]^T \in \mathbb{C}^N\), as well as the currents, flowing from say bus \(m\) to \(n\), denoted by \(I_{mn}\). Kirchoff’s law in vector-matrix form simply dictates \(\mathbf{v} = \mathbf{Y} \mathbf{v}\), where \(\mathbf{Y} \in \mathbb{C}^{N \times N}\) denotes the grid’s symmetric bus admittance matrix having \((m, n)\)-th entry given by

\[
Y_{mn} := \begin{cases} 
-y_{mn}, & \text{if } (m, n) \in \mathcal{E} \\
y_{mn} + \sum_{\nu \in \mathcal{N}_n} y_{\nu n}, & \text{if } m = n \\
0, & \text{otherwise} 
\end{cases}
\]  

(1)

with \(y_{mn}\) denoting the line admittance between buses \(m\) and \(n\); \(y_{mn}\) bus \(n\)’s admittance to the ground; and \(\mathcal{N}_n\) the set of all buses linked to bus \(n\) through transmission lines. In addition, the current flow is given by \(I_{mn} = y_{mn} V_m + y_{mn} (V_m - V_n)\), with \(y_{mn}\) standing for the shunt admittance at bus \(m\) associated with line \((m, n)\). Clearly, all current variables are linearly related to the state \(\mathbf{v}\). As for the nonlinear measurements, the AC power flow model asserts that the apparent power injection into bus \(n\) is given by \(P_n + jQ_n = V_n I_n^H\), while the apparent power flow from bus \(m\) to bus \(n\) by \(P_{mn} + jQ_{mn} = V_m I_m^H\). Further, expressing the squared bus voltage magnitude as \(|V_n|^2 = V_n I_n^H\), it is clear that all measurable quantities listed earlier are nonlinearly (in fact quadratically) related to \(\mathbf{v}\).

Apart from the nonlinearity present, another challenge present in the SE is due to grossly corrupted meter measurements (a.k.a. bad data). Statistical tests such as the largest normalized residuals of the weighted least-squares (WLS) estimation error are typically employed to reveal and remove bad data [16]. Alternatively, robust estimators, such as the least-absolute deviation, or Huber’s M-estimators have also been considered; see e.g., [1] Ch. 6. Motivated by recent advances in sparsity-aware robust statistical inference [9], the fresh look at robust SE (R-SE) advocated in this paper is by an overcomplete model for the outlying data. To this end, collect first the \(M\) measurements in the vector \(\mathbf{z} := \{P_n\}_{n \in \mathcal{N}}, \{Q_n\}_{n \in \mathcal{N}}, \{P_{mn}\}_{(m, n) \in \mathcal{E}}, \{Q_{mn}\}_{(m, n) \in \mathcal{E}}, \{|V_n|^2\}_{n \in \mathcal{N}}\), where the check mark differentiates measured values from the noise-free variable. Consider also the scalar variables \(\{a_{\ell}\}_{\ell = 1}^M\) one per measurement, taking the value \(a_{\ell} = 0\) if the \(\ell\)-th measurement obeys the nominal (outlier-free) model, and \(a_{\ell} \neq 0\) if it corresponds to a bad datum. This way, the nonlinear measurement model becomes

\[
z_{\ell} = h_{\ell}(\mathbf{v}) + \epsilon_{\ell} + a_{\ell}, \quad \ell = 1, \ldots, M
\]

(2)

where \(h_{\ell}(\cdot)\) captures the quadratic relationship specified by the aforementioned AC power flow equations, and the zero-mean additive Gaussian white noise (AWGN) \(\epsilon_{\ell}\) is assumed uncorrelated across meters with variance \(\sigma_{\epsilon}^2\).

Recovering both \(\mathbf{v}\) and the \(M \times 1\) vector \(\mathbf{a} := [a_1, \ldots, a_M]^T\) essentially reveals the state and identifies faulty measurements. However, the system in (2) with both \(\mathbf{v}\) and \(\mathbf{a}\) being unknown is under-determined, as the number of measurements \(M\) is always less than the number of unknowns \(N + M\). Instrumental to handling this under-determinacy will be the (arguably low) percentage of outliers, which gives rise to a (high) level of sparsity, that is the number of zero entries in \(\mathbf{a}\). The degree of sparsity will be further linked in the ensuing section with the notions of observability and identifiability of the outlier vector. By capitalizing on the sparsity of \(\mathbf{a}\), the goal of jointly estimating and identifying \(\mathbf{v}\) and \(\mathbf{a}\) can be achieved by the following outlier-sparsity-controlling criterion

\[
\{\hat{\mathbf{v}}, \hat{\mathbf{a}}\} := \arg \min_{\mathbf{v}, \mathbf{a}} \sum_{\ell = 1}^M w_{\ell} \left[|z_{\ell} - h_{\ell}(\mathbf{v}) - a_{\ell}|^2 + \lambda \|a\|_0\right]
\]

(3)

where \(w_{\ell} := 1/\sigma_{\epsilon}^2\) \(\forall \ell\), and \(\lambda > 0\) scales the regularization term which comprises the \(\ell_0\)-pseudonorm, i.e., the number of non-zero \(a_{\ell}\)’s that naturally controls the number of outliers in \(\mathbf{a}\). Even with linear models however, solving the optimization problem in (3) is NP-hard due to the \(\ell_0\)-norm regularization [4]. Before proposing efficient schemes for solving the under-determined problem in (3), the next section will provide observability and identifiability analysis to assess the ability of R-SE to cope with sparse outlier patterns.

III. OUTLIER OBSERVABILITY AND IDENTIFIABILITY

The goal of this section is to investigate fundamental uniqueness issues associated with the system under-determinacy arising due to the overcomplete outlier-aware model in (2). To isolate uniqueness from noise resilience issues, focus is placed on the noise-free outlier-aware measurement model written in vector form as

\[
\mathbf{z} = \mathbf{h}(\mathbf{v}) + \mathbf{a}
\]

(4)

\(^1\text{For consistency with other measurements, }|V_n|^2 \text{ is considered from now on. This is possible by adopting } |V_n|^2 = |V_n|^2 + \epsilon_V, \text{ where } \epsilon_V \text{ is zero-mean Gaussian with small variance } \sigma_{\epsilon_V}^2\).
with the high-dimensional function \( h(\cdot) : \mathbb{C}^N \rightarrow \mathbb{R}^M \).

**Definition 1.** Given measurements \( z = h(v_o) + a_o \), with \( v_o \) denoting the true state, and \( h(\cdot) \) known, the outlier vector \( a_o \) is identifiable if and only if (iff) \( \forall \ v_o \) the set

\[
\mathcal{V} := \{ v \in \mathbb{C}^N | \ z = h(v_o) + a_o = h(v) \}
\]

is empty. Furthermore, the outlier vector \( a_o \) is identifiable iff \( \forall \ v_o \) the set

\[
\mathcal{S} := \{ (v, a) | h(v) + a = z, \|a\|_0 \leq \|a_o\|_0 \}
\]

has only one element, namely \( (v_o, a_o) \).

But why are outlier observability and identifiability intuitively important? For an observable \( a_o \), upon collecting \( z \), the system operator can discern whether there are bad data or not. In addition, for an identifiable \( a_o \), the system operator can recover exactly (in the absence of nominal noise) both \( a_o \) and \( v_o \) in the presence of bad data.

Definition 1 implies that if \( a_o \) is identifiable, then it is necessarily observable, because otherwise the set \( \mathcal{V} \) in (5) would have at least one element \( v' \in \mathbb{C}^N \); in which case, the pair \( (v', 0) \) would be an additional second element of \( \mathcal{S} \) in (6) - a fact contradicting identifiability. Therefore, as a property of an outlier vector \( a_o \), identifiability is stronger than (i.e., subsumes) its observability.

Without accounting for the nominal AWGN in (2), it is possible to reduce the cost in (3) to only the \( \ell_0 \)-norm, while including the quadratic part as equality constraint to obtain

\[
\{ \tilde{v}, \tilde{a} \} := \arg \min_{h(v)+a=z} \|a\|_0
\]

\[
= \arg \min_{v, a=z-h(v)} \|z - h(v)\|_0. \quad (7)
\]

Clearly, for the noise-free R-SE problem in (7), the pair \( (v_o, a_o) \) is feasible, and the cost evaluated at \( (v_o, a_o) \) equals \( \|z - h(v_o)\|_0 = \|a_o\|_0 \). This is also the minimum cost attainable when \( a_o \) is identifiable, as there is no other pair \( (v, a) \) with smaller \( \ell_0 \)-norm \( \|a\|_0 \) according to Definition 1.

Conversely, if the noise-free problem (2) has a unique solution given by \( (v_o, a_o) \), then \( a_o \) is identifiable. Similarly, the noise-free R-SE formulation can easily detect the presence of bad data if the minimum achievable is non-zero. This clearly demonstrates the role of the outlier vector’s \( \ell_0 \)-norm in the R-SE criterion (3), in identifying the presence of bad data, or, in recovering the true state even when bad data are present.

A critical attribute for an observable (identifiable) outlier vector is its maximum sparsity level \( K_o \) (respectively \( K_i \)). To appreciate this, consider the two broad classes that outliers typically come from. The first class includes bad data emerging due to faulty meters, telemetry errors, or software bugs, which generally occur rarely, that is with low probability; see e.g., [9] and references therein. Here, \( K_o \) quantifies the maximum number of bad data that can be revealed with high probability; while \( K_i \) denotes the maximum number of outlying meters that can be identified so that recovery of the true state becomes feasible. The second source of outliers comprises malicious data attacks, in which the adversary can typically control only a subset of meters with limited cardinality [10, 13, 14]. In this class of outliers, \( K_o \) and \( K_i \) can suggest the minimum number of meters that must be protected to render malicious data attacks ineffective.

Even though \( K_o \) (\( K_i \)) is useful for assessing the degree of outlier observability (identifiability), deciding whether a given vector \( a_o \) is observable or identifiable for the nonlinear AC model (3) is challenging, except for the trivial case \( a_o = 0 \). Fortunately, it is possible to obtain \( K_o \) and \( K_i \) by leveraging the notion of the measurement distance for any nonlinear function \( h(\cdot) \), as defined next.

**Definition 2.** The measurement distance for the function \( h(\cdot) : \mathbb{C}^N \rightarrow \mathbb{R}^M \) is given by

\[
D(h) := \min_{v \neq v'} \| h(v) - h(v') \|_0
\]

\[
= \min_{v \neq v'} \sum_{\ell=1}^M \| h_\ell(v) - h_\ell(v') \|_0
\]

where \( \| \cdot \|_0 \) denotes the indicator function.

The notion of measurement distance parallels that of the Hamming distance in channel coding theory; see e.g., [8, Sec. 7.11]. Given any linear mapping over a known finite field, the Hamming distance characterizes the minimum difference between any two strings that lie in the mapped space, and it can be easily computed for fixed problem dimensions. However, for the R-SE problem of interest, \( v \) is drawn from the complex field \( \mathbb{C}^N \), while the mapping \( h(\cdot) \) is quadratic. Compared to the Hamming distance it will be generally very challenging to compute \( D(h) \) in (8).

Interestingly, as the Hamming distance has been popular due to its connection with the error control capability of linear channel codes, the measurement distance in (8) will turn out to be particularly handy in characterizing outlier observability and identifiability, as asserted in the following proposition.

**Proposition 1.** Given the measurement distance \( D \) for the nonlinear function \( h(\cdot) \) in (8), the maximum sparsity level of an observable outlier vector is \( K_o = D - 1 \), while the maximum one of an identifiable outlier vector is \( K_i = \lceil \frac{D-1}{2} \rceil \).

The proof for both statements follows readily from Definition 1 using simple contradiction arguments, and for this reason it is omitted. Notice that the second part can also be deduced after adapting [20, Thm. 5.1], which neither explicitly relates to the notion of measurement distance, nor it is linked with the maximum sparsity level of observable outliers.

Using the measurement distance metric, Proposition 1 provides a unifying framework to understand the tolerance of any function \( h(\cdot) \) to the number of outlying data. Since the measurement distance of any nonlinear function is difficult to obtain, the ensuing subsection pursues linearized approximants of the quadratic measurement model, which are typically employed by Gauss-Newton iterative SE solvers, and can be used to provide surrogate distance metrics. Depending on initialization, the linear approximants could not only be very
accurate, but will also shed light on understanding uniqueness issues associated with nonlinear AC power system models.

A. Linear Approximation Model

Consider linearizing the nonlinear measurement model [4] expressed in terms of the polar coordinates of the state vector, as in e.g., [11] Sec. 2.6. Toward this end, the \( N \times 1 \) complex vector \( \mathbf{v} \) is mapped first to the \( 2N \times 1 \) real vector \( \mathbf{x} := [|V_1|, \ldots, |V_N|, \angle V_1, \ldots, \angle V_N] \in \mathbb{R}^{2N} \). Invoking the first-order Taylor expansion, the noise-free \( \mathbf{z} \) can be approximated around a given point \( \tilde{\mathbf{v}} \), or the corresponding \( \tilde{\mathbf{x}} \), by

\[
\mathbf{z} = \mathbf{h}(\mathbf{v}) + \mathbf{a} \approx \mathbf{h}(\tilde{\mathbf{v}}) + \mathbf{H}_z (\mathbf{x} - \tilde{\mathbf{x}}) + \mathbf{a}
\]

where \( \mathbf{H}_z \in \mathbb{R}^{M \times (2N)} \) denotes the Jacobian matrix evaluated at \( \tilde{\mathbf{x}} \). Upon defining \( \tilde{\mathbf{z}} := \mathbf{z} - \mathbf{h}(\tilde{\mathbf{v}}) + \mathbf{H}_z \mathbf{x} \), the approximate model [2] becomes a linear one in the unknown \( \mathbf{x} \), that is

\[
\tilde{\mathbf{z}} \approx \mathbf{H}_z \mathbf{x} + \mathbf{a}.
\]

The measurement distance of the linear function in (10) can be found easily, as summarized next.

**Proposition 2.** The measurement distance for any linear mapping characterized by a full column-rank matrix \( \mathbf{H}_z \in \mathbb{R}^{M \times (2N)} \) is

\[
D = M + 1 - \text{rank}(\mathbf{H}_z).
\]

The proof relies on simple linear algebra arguments as follows. Using Definition 2, the measurement distance \( D := \min_{\mathbf{x} - \mathbf{x}' \neq 0} \| \mathbf{H}_z (\mathbf{x} - \mathbf{x}') \|_0 \) is attained when matrix \( \mathbf{H}_z \) has at most \( (M - D) \) linearly dependent rows; otherwise, the number of zero entries of \( \mathbf{H}_z (\mathbf{x} - \mathbf{x}') \) would be \( (M - D + 1) \) and that of non-zero ones \( (D - 1) \), which leads to a contradiction; hence, \( \text{rank}(\mathbf{H}_z) = M - D + 1 \), as asserted by Proposition 2.

Recalling from Proposition 1 how \( D \) is linked with the outlier observability and identifiability levels, the next corollary follows readily.

**Corollary 1.** For any linear mapping characterized by \( \mathbf{H}_z \), the maximum sparsity level of an observable outlier is \( K_o = M - \text{rank}(\mathbf{H}_z) \), while the maximum sparsity level of an identifiable outlier is \( K_i = \left\lfloor \frac{M - \text{rank}(\mathbf{H}_z)}{2} \right\rfloor \).

For the linear approximation model in (10), the measurement distance \( D \) grows linearly with the number of meters \( M \). This demonstrates that measurement redundancy is very beneficial for improving resilience to outliers. Conceivably, \( D \) could be further boosted thanks to the nonlinearity in \( \mathbf{h}(\cdot) \). Compared to its linear counterpart, the quadratic function \( \mathbf{h}(\cdot) \) is likely to increase the dimension of the space that is mapped to, and thus lead to a larger measurement distance in a space of higher dimensionality. This is precisely the reason why highly nonlinear functions find important applications to cryptography [7]. Although linearization provides a viable approximant, quantifying (or bounding) the measurement distance for the quadratic \( \mathbf{h}(\cdot) \) corresponding to the AC power flow model constitutes an interesting future research direction.

IV. SOLVING THE R-SE VIA SDR

This section will leverage convex relaxation techniques to solve the R-SE problem in (3). First, building on the premise of compressive sensing [6], the \( \ell_1 \)-norm can be employed to tackle the NP-hard \( \ell_0 \)-norm and relax the R-SE cost in (4) to

\[
\{ \hat{\mathbf{v}}, \hat{\mathbf{a}} \} := \arg \min_{\mathbf{v}, \mathbf{a}} \sum_{\ell=1}^M w_\ell |z_\ell - h_\ell(\mathbf{v}) - a_\ell|^2 + \lambda \| \mathbf{a} \|_1.
\]

The \( \ell_1 \)-norm relaxation has been used for various robust statistical inference tasks, with documented theoretical guarantees; see e.g., [9]. It has also been adopted recently for R-SE with the linear measurement model in [12], and also with the nonlinear one in [20] using the linearization technique employed by iterative Gauss-Newton SE solvers. However, the pertinent performance analysis in [20] has been given for a general nonlinear regression model, but not for the specific quadratic measurement model corresponding to AC power systems. Moreover, it is worth stressing that the iterative optimization framework adapted in [20] to solve (11) offers no guarantees regarding convergence or global optimality. In a nutshell, the desiderata remains to develop an R-SE solver capable of accounting for the practical AC quadratic measurement model, while attaining or approximating the global optimum at polynomial-time complexity.

This task will be pursued here using semidefinite relaxation (SDR), which has been recently recognized as a powerful technique for convexifying the SE with nonlinear measurement models [22]. To this end, each quadratic measurement \( z_\ell \) will be expressed linearly in terms of the outer-product matrix \( \mathbf{V} := \mathbf{v}\mathbf{v}^H \). Let \( \{ \mathbf{e}_n \}_{n=1}^N \) denote the canonical basis of \( \mathbb{R}^N \), and define the following admittance-related matrices

\[
\mathbf{Y}_n := \mathbf{e}_n\mathbf{e}_n^T, \quad Y_{mn} := \langle \mathbf{f}_{mn} + y_{mn} \rangle \mathbf{e}_m\mathbf{e}_m^T - y_{mn}\mathbf{e}_m\mathbf{e}_n^T
\]

and their related Hermitian counterparts

\[
\mathbf{H}_{P,n} := \frac{1}{2} (\mathbf{Y}_n + \mathbf{Y}_n^H), \quad \mathbf{H}_{Q,n} := \frac{j}{2} (\mathbf{Y}_n - \mathbf{Y}_n^H)
\]

\[
\mathbf{H}_{P,mn} := \frac{1}{2} (\mathbf{Y}_{mn} + \mathbf{Y}_{mn}^H), \quad \mathbf{H}_{Q,mn} := \frac{j}{2} (\mathbf{Y}_{mn} - \mathbf{Y}_{mn}^H)
\]

\[
\mathbf{H}_{V,n} := \mathbf{e}_n\mathbf{e}_n^T.
\]

Using these definitions, the following lemma is proved in [22] to establish a linear model in the complex rank-one matrix \( \mathbf{V} \).

**Lemma 1.** All error-free measurement variables are linearly related with the outer-product \( \mathbf{V} \) as

\[
P_n = \text{Tr}(\mathbf{H}_{P,n}\mathbf{V}), \quad Q_n = \text{Tr}(\mathbf{H}_{Q,n}\mathbf{V})
\]

\[
P_{mn} = \text{Tr}(\mathbf{H}_{P,mn}\mathbf{V}), \quad Q_{mn} = \text{Tr}(\mathbf{H}_{Q,mn}\mathbf{V})
\]

\[
|V_n|^2 = \text{Tr}(\mathbf{H}_{V,n}\mathbf{V}).
\]

Thus, the measurement \( z_\ell \) in (3) can be written as

\[
z_\ell = h_\ell(\mathbf{v}) + \epsilon_\ell + a_\ell = \text{Tr}(\mathbf{H}_z\mathbf{V}) + \epsilon_\ell + a_\ell
\]
where $H$ is a Hermitian matrix specified in accordance with \[13a\]–\[13c\].

Lemma 1 implies the following equivalent reformulation of \[11\] [cf. \[15\]]:

\[
\begin{align*}
\left\{ \hat{V}_1, \hat{a}_1 \right\} := \arg \min_{V, a} & \sum_{\ell=1}^M w_\ell [z_\ell - \text{Tr}(H_\ell V) - a_\ell]^2 + \lambda ||a||_1 \\
onumber
\text{s.to} & \quad V \in \mathbb{C}^{N \times N} \geq 0, \text{ and rank}(V) = 1
\end{align*}
\] \hfill (16a)

where the positive semi-definiteness and rank constraints jointly ensure that for any $V$ admissible to \[16b\], there always exists a state vector $v \in \mathbb{C}^N$ such that $V = vv^H$.

Albeit the linearity between $z_\ell$ and $V$ in the new formulation \[16\], nonconvexity is still present in two aspects: i) the cost in \[16a\] has degree 4 wrt the entries of $V$; and ii) the rank constraint in \[16b\] is nonconvex. Aiming for a semidefinite programming (SDP) formulation of \[16\], Schur’s complement lemma, see e.g., [5, Appx. 5.5], can be leveraged to convert the summands in \[16a\] to a linear cost over an auxiliary vector $\chi \in \mathbb{R}^M$. Specifically, with $w := [w_1, \ldots, w_M]^T$ and likewise for $\chi$, consider an R-SE reformulation as

\[
\begin{align*}
\left\{ \hat{V}_2, \hat{a}_2, \hat{\chi}_2 \right\} := \arg \min_{V, a, \chi} & \quad w^T \chi + \lambda ||a||_1 \\
\text{s.to} & \quad V \geq 0, \text{ and rank}(V) = 1, \\
& \quad \begin{bmatrix}
-\chi \ell & z_\ell - \text{Tr}(H_\ell V) - a_\ell \\
 z_\ell - \text{Tr}(H_\ell V) - a_\ell & -1
\end{bmatrix} \preceq 0 \forall \ell.
\end{align*}
\] \hfill (17a)

Upon adapting results from \[22\], the equivalence among all three R-SE formulations can be asserted as follows.

**Proposition 3.** For the AC power flow model, all three non-convex formulations in \[11\], \[16\], and \[17\], solve an equivalent R-SE problem. For the optima of these problems, it holds that

\[
\hat{V}_1 = \hat{V}_2 = \hat{v} \hat{v}^H \quad \text{and} \quad \hat{\chi}_{2, \ell} = \left(z_\ell - \text{Tr}(H_\ell \hat{V}_2)\right)^2 \forall \ell. \tag{18}
\]

Proposition 3 establishes the relevance of the novel R-SE formulation \[17\], which is still nonconvex though, due to the rank-1 constraint. Fortunately though \[17\] is amenable to the SDR technique, which amounts to dropping the rank constraint and has well-appreciated merits as an optimization tool; see e.g., [15] for a tutorial treatment of its applications in signal processing and communications. The contribution here consists in permeating the benefits of this powerful optimization tool to estimating the state of AC power systems, even when outliers (bad data or cyber-attacks) are present.

In the spirit of SDR, relaxing the rank constraint in \[17b\] leads to the following SDP formulation:

\[
\begin{align*}
\left\{ \hat{V}, \hat{a}, \hat{\chi} \right\} := \arg \min_{V, a, \chi} & \quad w^T \chi + \lambda ||a||_1 \\
\text{s.to} & \quad V \geq 0, \\
& \quad \begin{bmatrix}
-\chi \ell & z_\ell - \text{Tr}(H_\ell V) - a_\ell \\
z_\ell - \text{Tr}(H_\ell V) - a_\ell & -1
\end{bmatrix} \prec 0 \forall \ell.
\end{align*}
\] \hfill (19a)

SDR endows R-SE with a convex SDP formulation for which efficient schemes are available to obtain the global optimum using, e.g., the interior-point solver SeDuMi \[18\]. The worst-case complexity of this SDP solver is $O(M^4 \sqrt{N \log(1/\epsilon)})$ for a given solution accuracy $\epsilon > 0$ \[15\]. For typical power networks, $M$ is in the order of $N$, and thus the worst-case complexity becomes $O(N^{4.5} \log(1/\epsilon))$. Further computational complexity reduction is possible by exploiting the sparsity, and the so-called “chordal” data structure of matrix $V$, as detailed in \[22\].

Nonetheless, the SDP problem \[19\] is only a relaxed version of the equivalent R-SE in \[17\]; hence, its solution $V$ may have rank greater than 1, which makes it necessary to recover a feasible estimate $\hat{v}$ from $V$. This is possible by eigen-decomposing $V = \sum_{\ell=1}^M \lambda_\ell u_\ell u_\ell^H$, where $r := \text{rank}(V)$, $\lambda_1 \geq \cdots \geq \lambda_r > 0$ denotes the positive ordered eigenvalues, and $\{u_\ell \in \mathbb{C}^N\}_{\ell=1}^r$ are the corresponding eigenvectors. Since the best (in the minimum-norm sense) rank-one approximation of $V$ is $\lambda_1 u_1 u_1^H$, the state estimate can be chosen equal to $\hat{v}(u_1) := \sqrt{\lambda_1} u_1$.

Besides this eigenvector approach, **randomization** offers another way to extract an approximate R-SE vector from $V$, with quantifiable approximation accuracy; see e.g., [15]. The basic idea is to generate multiple Gaussian distributed random vectors $\nu \sim CN(0, V)$, and pick the one with the minimum error cost corresponding to the set of inlier meters $M_\ell := \{\ell | 1 \leq \ell \leq M_\ell, \hat{a}_\ell \neq 0\}$. Note that although any vector $\nu$ is feasible for \[11\], it is still possible to decrease the minimum achievable cost by rescaling to obtain $\hat{v}(\nu) = c\nu$, where the optimal weight can be chosen as the solution of the following convex problem as

\[
\hat{c} = \arg \min_{c > 0} \sum_{\ell \in M_\ell} w_\ell \left(z_\ell - c^2 \nu^H H_\ell \nu \right)^2
\]

\[
= \sqrt{\frac{\sum_{\ell \in M_\ell} w_\ell z_\ell \nu^H H_\ell \nu}{\sum_{\ell \in M_\ell} w_\ell \left(\nu^H H_\ell \nu \right)^2}} \tag{20}
\]

It will be of interest to find approximation bounds for the SDR-based R-SE approach, or, obtain meaningful conditions under which the relaxed solution coincides with the unrelaxed one. Both problems constitute interesting future directions for analytical research, while the ensuing section will demonstrate the performance improvement possible with the proposed method using numerical tests of practical systems.

**V. PRELIMINARY SIMULATIONS**

The novel SDR-based R-SE approach is tested in this section using the IEEE 30-bus system with 41 transmission lines \[17\], and compared to existing WLS methods that are based on Gauss-Newton iterations. The software toolbox MATPOWER \[21\] is used to generate the pertinent power flow and meter measurements. In addition, its SE function doSE has been adapted to realize the WLS Gauss-Newton iterations following from \[20\]. The iterations terminate either upon convergence, or, once the condition number of the approximate linearization exceeds $10^8$, which flags divergence of the iterates. To solve
Further enhancements to the SDR-based R-SE framework are currently pursued toward developing more efficient and tailored solvers by exploiting sparsity of the SDP problem structure. The measurement distance for the power meter quadratic functions is also under investigation using insights from nonlinear channel coding theory.

VI. CONCLUSIONS AND CURRENT RESEARCH

For the practical nonlinear AC power system model, uniqueness issues and robust state estimation (R-SE) algorithms were investigated in this paper, when outliers (bad data and/or malicious attacks) are present. Using a sparse overcomplete outlier model, observability and identifiability issues were quantified using the notion of measurement distance for the quadratic measurement model. Valuable insights and computable levels of outlier observability and identifiability were provided for linear approximations of the quadratically nonlinear models. A novel SDR-based scheme was also developed by tactfully relaxing the nonconvex R-SE problem to a convex SDP one, thus rendering it efficiently solvable via existing interior-point methods. Preliminary numerical simulations on the 30-bus benchmark system demonstrated improved performance of the proposed R-SE scheme.

REFERENCES

[1] A. Abur and A. G. Exposito, Power System State Estimation: Theory and Implementation, New York, NY: Marcel Dekker, 2004.
[2] A. R. Bergen and V. Vittal, Power System Analysis, 2nd ed., Upper Saddle River, N.J.: Prentice Hall, 2000.
[3] D. P. Bertsekas, Nonlinear Programming, 2nd ed., Belmont, MA: Athena Scientific, 1995.
[4] R. B. Bobba, K. M. Rogers, Q. Wang, H. Khurana, K. Najmeddine, and T. J. Overbye, “Detecting false data injection attacks on dc state estimation,” Proc. First Workshop on Secure Control Systems, Stockholm, Sweden, Apr. 2010.
[5] S. Boyd and L. Vandenberghe, Convex Optimization, Cambridge, UK: Cambridge University Press, 2004.
[6] E. J. Candès and T. Tao, “Near-optimal signal recovery from random projections: Universal encoding strategies?” IEEE Trans on Info. Theory, vol. 52, pp. 5406–5425, Dec. 2006.
[7] C. Carlet and C. Ding, “Highly nonlinear mappings,” Journal of Complexity, vol. 20, no. 2-3, pp. 205-244, Apr. 2004.
[8] T. M. Cover and J. A. Thomas, Elements of Information Theory, 2nd ed., NJ: John Wiley & Sons, 2006.
[9] G. B. Giannakis, G. Mateos, S. Farahmand, and H. Zhu, “USPACOR: Universal sparsity-controlling outlier rejection,” Proc. of Intl. Conf. on Acoustics, Speech and Signal Processing, Prague, Czech Republic, May 22–27, 2011.
[10] A. Gomez-Exposito, A. Abur, P. Rousseaux, A. de la Villa Jaen, and C. Gomez-Quiles, “On the use of PMUs in power system state estimation,” Proc. 17th Power Systems Computation Conference, Stockholm, Sweden, Aug. 22-26, 2011.
[11] M. Grant and S. Boyd, “CVX: Matlab software for disciplined convex programming,” version 1.21, Apr. 2011. [Online]. Available: http://cvxr.com/cvx/
[12] V. Kekatos and G. B. Giannakis, “Distributed robust power system state estimation,” IEEE Transactions on Power Systems, 2012 (submitted). [Online]. Available: http://arxiv.org/abs/1204.0991
[13] O. Kosut, L. Jia, J. Thomas, and L. Tong, “Malicious data attacks on the smart grid,” IEEE Trans. Smart Grid, vol. 2, no. 4, pp. 645–658, Dec. 2011.
[14] Y. Liu, M. K. Reiter, and P. Ning, “False data injection attacks against state estimation in electric power grids,” Proc. ACM Conf. on Computer and Comm. Security, Chicago, IL, Nov. 2009, pp. 21–32.
[15] Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, and S. Zhang, “Semidefinite relaxation of quadratic optimization problems,” IEEE Signal Processing Magazine, vol. 27, no. 3, pp. 20–34, May 2010.
[16] A. Monticelli, “Electric power system state estimation,” Proc. of the IEEE, vol. 88, no. 2, pp. 262–282, Feb. 2000.
[17] Power Systems Test Case Archive, University of Washington. [Online] Available: http://www.ee.washington.edu/research/pstca/.
[18] J. F. Sturm, “Using SeDuMi 1.02, a Matlab toolbox for optimization over symmetric cones,” Optimization Methods Software, vol. 11–12, pp. 625–653, Aug. 1999. [Online]. Available: http://sedumi.mcmaster.ca
[19] U.S. Department of Energy, The Smart Grid: An introduction, 2008. [Online.] Available: http://www.oe.energy.gov/SmartGridIntroduction.htm
[20] W. Xu, M. Wang, and A. Tang, “Sparse recovery from nonlinear measurements with applications in bad data detection for power networks,” Dec. 2011 (submitted). [Online]. Available: http://arxiv.org/abs/1112.6234
[21] R. D. Zimmerman, C. E. Murillo-Sanchez, and R. J. Thomas, “Matpower: Steady-state operations, planning and analysis tools for power systems research and education,” IEEE Trans. on Power Systems, vol. 26, no. 1, pp. 12–19, Feb. 2011.
[22] H. Zhu and G. B. Giannakis, “Estimating the state of AC power systems using semidefinite programming,” Proc. the 43rd NAPS, pp. 1-7, Boston, MA, Aug. 4-6, 2011.