Sample Complexity of Power System State Estimation using Matrix Completion

Preprint

Joshua Comden[1], Marcello Colombinoy[2], Andrey Bernstein[3], and Zhenhua Liu[1]

1 Stony Brook University
2 McGill University
3 National Renewable Energy Laboratory

Presented at the IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids
Beijing, China
October 21-24, 2019

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Contract No. DE-AC36-08GO28308

Conference Paper
NREL/CP-5D00-74580
November 2019

Sample Complexity of Power System State Estimation using Matrix Completion

Preprint

Joshua Comden[1], Marcello Colombinoy[2], Andrey Bernstein[3], and Zhenhua Liu[1]

1 Stony Brook University
2 McGill University
3 National Renewable Energy Laboratory

Suggested Citation
Comden, Joshua, Marcello Colombinoy, Andrey Bernstein, and Zhenhua Liu. 2019. *Sample Complexity of Power System State Estimation using Matrix Completion: Preprint*. Golden, CO: National Renewable Energy Laboratory. NREL/ CP-5D00-74580. https://www.nrel.gov/docs/fy20osti/74580.pdf.

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Contract No. DE-AC36-08GO28308

Conference Paper
NREL/CP-5D00-74580
November 2019

National Renewable Energy Laboratory
15013 Denver West Parkway
Golden, CO 80401
303-275-3000 • www.nrel.gov

NOTICE

This work was authored in part by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. This work was supported by the Laboratory Directed Research and Development (LDRD) Program at NREL. The views expressed herein do not necessarily represent the views of the DOE or the U.S. Government.

Cover Photos by Dennis Schroeder: (clockwise, left to right) NREL 51934, NREL 45897, NREL 42160, NREL 45891, NREL 48097, NREL 46526.

NREL prints on paper that contains recycled content.

Sample Complexity of Power System State Estimation using Matrix Completion

Joshua Comden*, Marcello Colombino†, Andrey Bernstein‡, and Zhenhua Liu*
*Stony Brook University, {joshua.comden, zhenhua.liu}@stonybrook.edu
†McGill University, marcello.colombino@mcgill.ca
‡National Renewable Energy Laboratory, andrey.bernstein@nrel.gov

Abstract—**In this paper, we propose an analytical framework to quantify the amount of data samples needed to obtain accurate state estimation in a power system — a problem known as *sample complexity* analysis in computer science. Motivated by the increasing adoption of distributed energy resources into the distribution-level grids, it becomes imperative to estimate the state of distribution grids in order to ensure stable operation. Traditional power system state estimation techniques mainly focus on the transmission network which involve solving an overdetermined system and eliminating bad data. However, distribution networks are typically *underdetermined* due to the large number of connection points and high cost of pervasive installation of measurement devices. In this paper, we consider the recently proposed state-estimation method for underdetermined systems that is based on *matrix completion*. In particular, a *constrained* matrix completion algorithm was proposed, wherein the standard matrix completion problem is augmented with additional equality constraints representing the physics (namely power-flow constraints). We analyze the sample complexity of this general method by proving an upper bound on the sample complexity that depends directly on the properties of these constraints that can lower number of needed samples as compared to the unconstrained problem. To demonstrate the improvement that the constraints add to state estimation, we test the method on a 141-bus distribution network case study and compare it to the traditional least squares minimization state estimation method.**

I. Introduction

State estimation is one of the fundamental data analysis tasks in power systems. In its classical form, it amounts to estimating voltage phasors at all the buses of the network given some data gathered from the network [5]. It has a long and established history in transmission networks, where classical approaches based on weighted least-squares methods are applicable due to full observability of the network [6]. The latter conditions roughly speaking mean that the underlying system of equations for the estimation problem is *overdetermined*, i.e., it has more observables (and, hence, equations) than unknown variables [7]. In traditional distribution networks however, state estimation is typically not used, or used very rarely [8]. Unlike in transmission networks, there is a lack of pervasive installation of measurement devices such as phasor measurement units (PMUs) [9], [10]. Hence, the estimation problem is *underdetermined* and so classical, simple approaches (e.g. weighted least-squares) cannot be applied since they require full observability [11].

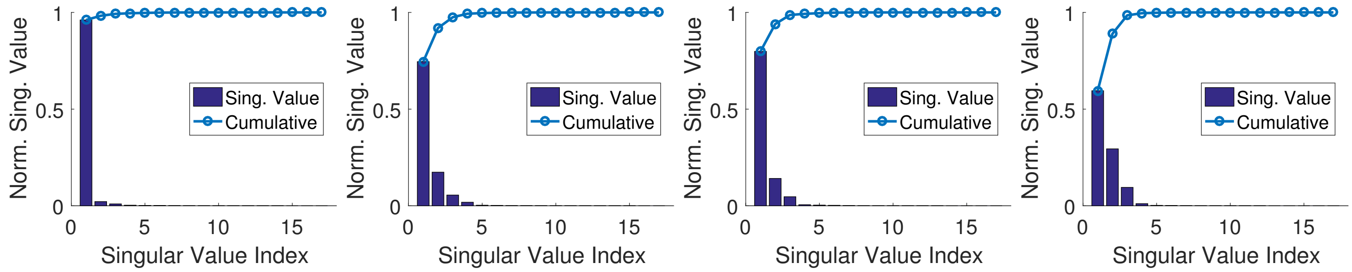
However, recently, distribution networks have undergone a radical change due to massive penetration of distributed energy resources (DERs) at the edge of the network [12], [13], [14], [15]. This creates both challenges and opportunities. On the challenges side, DERs (and especially renewable energy resources such as photovoltaic panels and wind farms) introduce a lot of uncertainty into the system [16], [17], [18], [19]. Thus, accurate real-time state estimation is needed to ensure stable and safe operation of the network [20], [21]. On the opportunities side, the vast deployment of DERs introduces both control and measurement points that now allow the application of modern machine learning and data analytics methods to deal with problems such as state estimation [22], [23]. However, observability is still an issue: the corresponding estimation problem is typically underdetermined.

In this paper, we consider the recently proposed method for state estimation in underdetermined systems using *low-rank matrix completion* [24]. The method is based on augmenting the standard matrix completion approach [25] with power-flow constraints which provide an additional link between parameter values. As shown in [24] numerically with extensive simulations, this structured (or *physics-based*) approach performs very well in distribution networks under realistic low-observability scenarios. In the present paper, we set our goal to study the sample complexity of this approach.

Roughly speaking, sample complexity is the amount of data samples needed to obtain accurate estimation of the true state. Sample complexity in power-system state estimation is largely unexplored; even in the case of the classical weighted least-squares methods, the literature is scarce on the topic, whereas there is active research in computer science and machine learning community on the topic [26], [27]. However, the results of the standard matrix completion problem [25], [28] are somewhat conservative for direct application to distribution networks because the state of a power system is constrained by well known physical laws and cannot just be any set of random values. The main theoretical challenge is on how to measure the additional information coming from the physical constraints in terms of the amount of sampled state variables. The information can then be used to partially replace the need to make a specific number of measurements.

Therefore, this paper makes the following contributions:

1) We model the distribution network state estimation problem as a low-rank matrix completion problem and incorporate its physical constraints (Section II).



(a) 33 Bus [1] (b) 69 Bus [2] (c) 85 Bus [3] (d) 141 Bus [4]

Fig. 1. Singular values of matrices that represent the states of four different IEEE radial distribution test cases. All singular values are normalized by their sum. The bars are the individual values representing the fraction of additional information added by that rank-1 matrix, while the circles are the cumulative values representing the fraction of information held by it and the previous rank-1 matrices. In all cases, more than 95% of the state information can be recovered by just 3 rank-1 matrices. From this evidence, it shows that low-rank matrices can be used to accurately approximate the state of distribution networks.

2) We incorporate constraints into the sample complexity analysis of the standard (unconstrained) matrix completion to obtain a lower theoretical bound (Section III).
3) We verify the significant reduction in necessary sample sizes through real-world data based numerical evaluations using a distribution network test case (Section IV).

To the best of our knowledge, these are the first results in the literature on sample complexity for constrained matrix completion in general, and for state estimation on power systems in particular.

II. Problem Formulation

Notation: A column vector $\mathbf{x}$ is represented by a bold lowercase letter and a matrix $\mathbf{X}$ is represented by bold uppercase letter, while a scalar x or an entry X_{ij} are not bold and can be either upper or lower case. For complex number x, let $\text{Re}(x)$, $\text{Im}(x)$, and $|x|$ be its real component, its imaginary component, and its magnitude respectively. The kth matrix $\mathbf{X}^{(k)}$ in a sequence may be labeled by a superscript in parenthesis. $\mathbf{I}_n$ is the $n \times n$ identity matrix. A calligraphic letter $\mathcal{X}$ can be a set, vector space, or operator which will be distinctly made clear in context. Specifically, $\mathcal{P}_{\mathcal{X}}$ is the orthogonal projection onto vector space $\mathcal{X}$. The perpendicular vector space to $\mathcal{X}$ is $\mathcal{X}^{\perp}$. The transpose of matrix $\mathbf{X}$ is $\mathbf{X}^{\intercal}$. The ℓ_2-norm of vector $\mathbf{x}$ is $\|\mathbf{x}\|$. The Euclidean inner product of matrices $\mathbf{A} \in \mathbb{R}^{n_1 \times n_2}$ and $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2}$ is $\langle \mathbf{A}, \mathbf{X} \rangle := \text{trace}(\mathbf{A}^{\intercal}\mathbf{X})$. The Frobenius norm of matrix $\mathbf{X}$ is $\|\mathbf{X}\|_F := \sqrt{\langle \mathbf{X}, \mathbf{X} \rangle} = \sqrt{\sum_i \sum_j |X_{ij}|^2}$. The nuclear norm of matrix $\mathbf{X}$ is denoted by $\|\mathbf{X}\|_*$ and is the sum of the its singular values, while the spectral norm is denoted $\|\mathbf{X}\|$ and is the value of its largest singular value.

A. Power System Model

Consider a power network with n_{b} PQ buses in the set $\mathcal{N}$ which are buses that have set real and reactive power injections, and n_{l} lines in the set $\mathcal{L} \subset \mathcal{N} \times \mathcal{N}$. For each line $(s,t) \in \mathcal{L}$, bus s is denoted as the "From" bus and bus t is denoted as the "To" bus. Typically in a radial distribution network, the slack or feeder bus is labeled as bus 1 and all other buses are labeled sequentially outward so that when the lines are directed away from the feeder, the From bus has a smaller index.

Complex power is split into its real and reactive components represented by P and Q respectively. Power flows across lines are treated as injections into the line from both the From and To sides so that their sum equals the power loss:

$$P_{s,t}^{\text{From}} + P_{s,t}^{\text{To}} = P_{s,t}^{\text{Loss}} : \qquad \forall (s,t) \in \mathcal{L} \tag{1a}$$

$$Q_{s,t}^{\text{From}} + Q_{s,t}^{\text{To}} = Q_{s,t}^{\text{Loss}} : \qquad \forall (s,t) \in \mathcal{L}. \tag{1b}$$

Therefore from the conservation of power at each bus, its power injection into the bus must equal the power injections into the lines it is connected to:

$$P_s = \sum_{t:(s,t)\in\mathcal{L}} P_{s,t}^{\text{From}} + \sum_{t:(t,s)\in\mathcal{L}} P_{t,s}^{\text{To}} : \qquad \forall s \in \mathcal{N} \tag{2a}$$

$$Q_s = \sum_{t:(s,t)\in\mathcal{L}} Q_{s,t}^{\text{From}} + \sum_{t:(t,s)\in\mathcal{L}} Q_{t,s}^{\text{To}} : \qquad \forall s \in \mathcal{N}. \tag{2b}$$

The complex current injection I_s at each bus s and the complex current flow $I_{s,t}$ across each line (s,t) follow Kirchhoff's Current Law:

$$\sum_{t:(s,t)\in\mathcal{L}} \text{Re}\left(I_{s,t}\right) = \text{Re}\left(I_s\right) + \sum_{t:(t,s)\in\mathcal{L}} \text{Re}\left(I_{t,s}\right) : \quad \forall s \in \mathcal{N} \tag{3a}$$

$$\sum_{t:(s,t)\in\mathcal{L}} \text{Im}\left(I_{s,t}\right) = \text{Im}\left(I_s\right) + \sum_{t:(t,s)\in\mathcal{L}} \text{Im}\left(I_{t,s}\right) : \quad \forall s \in \mathcal{N}. \tag{3b}$$

Additionally, using the complex voltage V_s at each bus, Ohm's Law relates the voltage difference between the two sides of a line to its current flow:

$$\text{Re}\left(I_{s,t}\right) = G_{s,t}(\text{Re}(V_s) - \text{Re}(V_t)) - B_{s,t}(\text{Im}(V_s) - \text{Im}(V_t)), \quad \forall (s,t) \in \mathcal{L} \tag{4a}$$

$$\text{Im}\left(I_{s,t}\right) = B_{s,t}(\text{Re}(V_s) - \text{Re}(V_t)) + G_{s,t}(\text{Im}(V_s) - \text{Im}(V_t)), \quad \forall (s,t) \in \mathcal{L} \tag{4b}$$

where $G_{s,t}$ and $B_{s,t}$ are the conductance and susceptance of line (s,t). The power injections into each line from either side are determined from its current flow and voltage on that side:

$$P_{s,t}^{\text{From}} = \text{Re}(V_s)\text{Re}(I_{s,t}) + \text{Im}(V_s)\text{Im}(I_{s,t}) : \forall (s,t) \in \mathcal{L} \tag{5a}$$

$$Q_{s,t}^{\text{From}} = \text{Im}(V_s)\text{Re}(I_{s,t}) - \text{Re}(V_s)\text{Im}(I_{s,t}) : \forall (s,t) \in \mathcal{L} \tag{5b}$$

$$P^{\text{To}}_{s,t} = -(\text{Re}(V_t)\text{Re}(I_{s,t}) + \text{Im}(V_t)\text{Im}(I_{s,t})) : \forall (s,t) \in \mathcal{L} \quad (5c)$$

$$Q^{\text{To}}_{s,t} = -(\text{Im}(V_t)\text{Re}(I_{s,t}) - \text{Re}(V_t)\text{Im}(I_{s,t})) : \forall (s,t) \in \mathcal{L} \quad (5d)$$

The power loss across each line are determined from its magnitude of the current flow $|I_{s,t}|$:

$$P^{\text{Loss}}_{s,t} = R_{s,t}|I_{s,t}|^2 : \qquad \forall (s,t) \in \mathcal{L} \quad (6a)$$

$$Q^{\text{Loss}}_{s,t} = X_{s,t}|I_{s,t}|^2 : \qquad \forall (s,t) \in \mathcal{L} \quad (6b)$$

where $R_{s,t}$ and $X_{s,t}$ are the resistance and reactance of line (s,t), respectively.

Trivially, we also have the magnitudes of the complex voltages and currents derived from their real and imaginary parts:

$$|V_s| = \sqrt{\text{Re}(V_s)^2 + \text{Im}(V_s)^2} \qquad \forall s \in \mathcal{N} \quad (7a)$$

$$|I_s| = \sqrt{\text{Re}(I_s)^2 + \text{Im}(I_s)^2} \qquad \forall s \in \mathcal{N} \quad (7b)$$

$$|I_{s,t}| = \sqrt{\text{Re}(I_{s,t})^2 + \text{Im}(I_{s,t})^2} \qquad \forall (s,t) \in \mathcal{L}. \quad (7c)$$

B. State Estimation Problem

We represent the state of the power system in a block matrix $\mathbf{M}$ where one matrix $\mathbf{M}_{\text{b}}$ holds the state of the buses and the other matrix $\mathbf{M}_{\text{l}}$ holds the state of the lines

$$\mathbf{M} := \left[\begin{array}{c|c} \mathbf{M}_{\text{b}} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{M}_{\text{l}} \end{array}\right].$$

The state of the buses $\mathbf{M}_{\text{b}}$ is in an $n_{\text{b}} \times 8$ matrix which holds the following values in the row associated with bus $s \in \mathcal{N}$:

$$(P_s, Q_s, \text{Re}(V_s), \text{Im}(V_s), |V_s|, \text{Re}(I_s), \text{Im}(I_s), |I_s|)$$

while the state of the lines $\mathbf{M}_{\text{l}}$ is in an $n_{\text{l}} \times 9$ matrix which holds the following values in the row associated with line $(s,t) \in \mathcal{L}$:

$$\big(P^{\text{From}}_{s,t}, Q^{\text{From}}_{s,t}, P^{\text{To}}_{s,t}, Q^{\text{To}}_{s,t}, P^{\text{Loss}}_{s,t}, Q^{\text{Loss}}_{s,t}, \text{Re}(I_{s,t}), \text{Im}(I_{s,t}), |I_{s,t}|\big).$$

Classically, the state is represented in a more compact form of only the complex bus voltages since, given voltages, all other variables can be computed using (1)-(7). However, complex voltages can only be measured by PMUs which are expensive and are not available at almost all of the buses or lines in a distribution network. On the other hand, measurements of some of the other variables are more widely available such as $(P_s, Q_s, |V_s|, |I_s|)$ for a bus and $\big(P^{\text{From}}_{s,t}, Q^{\text{From}}_{s,t}, P^{\text{To}}_{s,t}, Q^{\text{To}}_{s,t}, |I_{s,t}|\big)$ for a line. Let Ω be the set of state matrix locations which have available measurements.

Therefore, the goal of this *under-determined* state estimation problem is to accurately fill in any unmeasured values in the state matrix $\mathbf{M}$, especially the complex bus voltages, using the available measured values from locations Ω and the power system equations (1)-(7). To that end, it was recently proposed in [24] to leverage the approximately low-rank structure of the state matrix (as demonstrated in Figure 1) to find a *minimum rank matrix* that satisfies (1)-(7) and matches the measured state values:

$$\begin{aligned} \min_{\mathbf{X}} \quad & \text{rank}(\mathbf{X}) \\ \text{s.t.} \quad & X_{ij} = M_{ij} \quad \forall (i,j) \in \Omega \\ & (1) - (7). \end{aligned}$$

However, there are two issues with the above problem formulation that make it non-convex, thus computationally hard to solve: (i) the objective function is non-convex; and (ii) the equality constraints (5)-(7) are not linear, therefore the feasible solution space for $\mathbf{X}$ is non-convex. To tackle the first challenge, a standard relaxation using *nuclear norm* [25] is used. To tackle the second challenge, constraints (5)-(7) are replaced with their *linear approximation*. The relationship of the voltage magnitude difference across a line and complex power flow $\big(P^{\text{Flow}}_{s,t}, Q^{\text{Flow}}_{s,t}\big)$ on that line can be linearly approximated for a radial distribution network [29]

$$|V_t| - |V_s| = \frac{1}{|V_1|}\big(R_{s,t}P^{\text{Flow}}_{s,t} + X_{s,t}Q^{\text{Flow}}_{s,t}\big) \quad \forall (s,t) \in \mathcal{L} \quad (9)$$

where V_1 is the voltage of the slack bus. It assumes either that the lines have no losses or that the power flow is so low that losses are negligible. Since the state $\mathbf{M}$ does not make this assumption and does not encode the power flows directly, we can approximate the power flow by taking the average power injection into the line, i.e. $P^{\text{Flow}}_{s,t} := \big(P^{\text{From}}_{s,t} - P^{\text{To}}_{s,t}\big)/2$ and $Q^{\text{Flow}}_{s,t} := \big(Q^{\text{From}}_{s,t} - Q^{\text{To}}_{s,t}\big)/2$.

C. Constrained Matrix Completion and Sample Complexity

The under-determined state estimation problem with linear system equations can be generalized to the following low-rank matrix completion problem with h linear equality constraints:

$$\min_{\mathbf{X}} \quad \text{rank}(\mathbf{X}) \quad (10a)$$

$$\text{s.t.} \quad X_{ij} = M_{ij} \qquad \forall (i,j) \in \Omega \quad (10b)$$

$$\langle \mathbf{A}^{(l)}, \mathbf{X} \rangle = b^{(l)} \qquad \forall l \in \{1, \dots, h\} \quad (10c)$$

where the matrix inner product is defined as $\langle \mathbf{A}, \mathbf{X} \rangle := \text{trace}(\mathbf{A}^\intercal \mathbf{X})$. Let m be cardinality of Ω and assume that the locations of $\mathbf{M}$ that make up Ω are sampled uniformly at random. The question for this general constrained matrix completion problem becomes *how large does m need to be so that the solution to Problem* (10) *is guaranteed to exactly match* $\mathbf{M}$*?* This value of m is is referred to as *sample complexity*.

Notice that if a matrix has rank r, then it also means that it has r nonzero singular values. Therefore, a simple heuristic of minimizing the sum of its singular values is used to approximate the minimization its rank [30]. This heuristic is actually the definition of the nuclear norm which is convex:

$$\|\mathbf{X}\|_* := \sum_{k=1}^{r} \sigma_k(\mathbf{X})$$

where $\sigma_k(\mathbf{X})$ is the kth largest singular value. With the substitution of the nuclear norm in place of the rank operator, we reformulate the matrix completion problem (10) to be

$$\min_{\mathbf{X}} \quad \|\mathbf{X}\|_* \tag{11a}$$
$$\text{s.t.} \quad X_{ij} = M_{ij} \qquad \forall (i,j) \in \Omega \tag{11b}$$
$$\langle \mathbf{A}^{(l)}, \mathbf{X} \rangle = b^{(l)} \qquad \forall l \in \{1, \ldots, h\} \tag{11c}$$

with the same question as before on the sample complexity for Ω under uniform random sampling.

III. Main Result on Sample Complexity

In this section, we formulate a result on sample complexity that takes advantage of the linear equality constraints in the problem formulation. The main challenge is on how to measure the information from the added constraints in terms of sample size which can be used to partially replace the need for extra measurements. The intuition behind the usefulness of the added constraints (10c) is that each constraint may eliminate a single degree of freedom from the feasible solution set. Thus, a set of constraints may decrease the search space for an approximation method so that less samples are needed to recover the underlying matrix $\mathbf{M}$.

Let $\mathbf{M}$ be an $n_1 \times n_2$ matrix of rank r which satisfies:

$$\mathbf{M} = \sum_{k=1}^{r} \sigma_k \mathbf{u}_k \mathbf{v}_k^{\intercal} \tag{12a}$$
$$\langle \mathbf{A}^{(l)}, \mathbf{M} \rangle = b^{(l)} \quad \forall l \in \{1, \ldots, h\} \tag{12b}$$

where (12a) is its Singular Value Decomposition (SVD). Without loss of generality, we assume that $n_1 \geq n_2$. The vectors $\mathbf{u}_1, \ldots, \mathbf{u}_r$ are unit vectors of size n_1 that are orthogonal to each other and the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are unit vectors of size n_2 that are orthogonal to each other. The scalars $\sigma_1, \ldots, \sigma_r$ which are used to linearly combine the matrices $\mathbf{u}_1\mathbf{v}_1^{\intercal}, \ldots, \mathbf{u}_r\mathbf{v}_r^{\intercal}$ to be equal to $\mathbf{M}$ are called its singular values. By convention, the singular values are listed in decreasing order so that σ_k refers to the kth largest singular value in (12a). The number of degrees of freedom of any $n_1 \times n_2$ matrix of rank r is $r(n_1 + n_2 - r)$.

A. High-Probability Exact Completion

Due to the probabilistic nature of the question on sample complexity, the answer will also be probabilistic. This is because for any given number of samples taken that is less than $(n_1-1)n_2$, there is some probability that the sampled locations will miss an entire row and thus have no information that can be used to recover it. Thus, our goal will be to determine how large does m, the cardinality of Ω, need to be to ensure a high probability of exact completion using the optimal solution to Problem (11). Another way to frame the objective is to find the conditions on m and $\mathbf{M}$ such that $\mathbf{M}$ is the unique solution to (11) with some probability.

A property of the underlying matrix $\mathbf{M}$ that must be understood is how well its information is spread among its columns and rows. A matrix with its information not well spread will require many samples. For this reason, [25] defines a property on the space spanned by either $(\mathbf{u}_1, \ldots, \mathbf{u}_r)$ or $(\mathbf{v}_1, \ldots, \mathbf{v}_r)$ which measures the spread of the weight of its elements compared to the standard basis, called *coherence*.

Definition 1. *For any subspace $\mathcal{U}$ in $\mathbb{R}^n$ with dimension r, let the coherence of $\mathcal{U}$ be defined as*

$$\mu(\mathcal{U}) := \frac{n}{r} \max_{i \in \{1,\ldots,n\}} \|\mathbf{P}_{\mathcal{U}} \mathbf{e}_i\|^2$$

where $\mathbf{P}_{\mathcal{U}}$ is the orthogonal projection matrix onto $\mathcal{U}$ and $\mathbf{e}_i$ is the i-th standard basis vector with dimension n.

With the following assumption, the lack of spread of information within $\mathbf{M}$ can be bounded by bounding the coherence of the spaces defined by the vectors in its SVD (12a).

Assumption 1. *The coherence of $\mathcal{U} :=$ span$(\mathbf{u}_1, \ldots, \mathbf{u}_r)$ and the coherence of $\mathcal{V} :=$ span$(\mathbf{v}_1, \ldots, \mathbf{v}_r)$ are both upper bounded by some constant $\mu_0 > 0$, i.e.*

$$\max\{\mu(\mathcal{U}), \mu(\mathcal{V})\} \leq \mu_0$$

To limit the concentration of information in the subgradient of the nuclear norm at $\mathbf{M}$ for any specific matrix location, an assumption is placed on the maximum value of sum of the rank-1 matrices $\mathbf{u}_1\mathbf{v}_1^{\intercal}, \ldots, \mathbf{u}_r\mathbf{v}_r^{\intercal}$ through the parameter ν_0.

Assumption 2. *The absolute value of each element in $\sum_{k=1}^{r} \mathbf{u}_k \mathbf{v}_k^{\intercal}$ is upper bounded by $\nu_0\sqrt{r/(n_1 n_2)}$ for some constant ν_0.*

One important item needed in proving that $\mathbf{M}$ is the unique solution to (11) is a vector space of matrices $\mathcal{T}$ that contains all $n_1 \times n_2$ matrices which have a column space in $\mathcal{U} :=$ span$(\mathbf{u}_1, \ldots, \mathbf{u}_r)$, i.e. the column space of $\mathbf{M}$, and all $n_1 \times n_2$ matrices which have a row space in $\mathcal{V} :=$ span$(\mathbf{v}_1, \ldots, \mathbf{v}_r)$, i.e. the row space of $\mathbf{M}$. Specifically, a vector space $\mathcal{T}$ of matrices is built from all the combinations of $\mathbf{u}_1, \ldots, \mathbf{u}_r$ that can span the column space and all the combinations of $\mathbf{v}_1, \ldots, \mathbf{v}_r$ that can span the row space via their outer products with the vectors $\{\mathbf{x}_1, \ldots, \mathbf{x}_r\} \in \mathbb{R}^{n_2}$ and $\{\mathbf{y}_1, \ldots, \mathbf{y}_r\} \in \mathbb{R}^{n_1}$:

$$\mathcal{T} := \left\{ \sum_{k=1}^{r} (\mathbf{u}_k \mathbf{x}_k^{\intercal} + \mathbf{y}_k \mathbf{v}_k^{\intercal}) : \mathbf{x}_k \in \mathbb{R}^{n_2}, \mathbf{y}_k \in \mathbb{R}^{n_1} \right\}$$

This vector space has a dimension of $r(n_1 + n_2 - r)$ which is equal to the degrees of freedom in any $n_1 \times n_2$ matrix of rank r.

To measure the amount of useful information held in the linear equalities (12b) that can explain $\mathbf{M}$, we develop quantities similar to the upper bounds of Assumptions 1 and 2 for the vector space spanned by $\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(h)}$, denoted by $\mathcal{Q}$. First, we measure how much of the vector space $\mathcal{T}$ remains uncovered by $\mathcal{Q}$:

$$\mu_{\mathcal{Q}^\perp} := \frac{\sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \left\| \mathcal{P}_{\mathcal{T}} \mathcal{P}_{\mathcal{Q}^\perp} \left(\mathbf{e}_i \mathbf{e}_j^{\intercal} \right) \right\|_F^2}{\sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \left\| \mathcal{P}_{\mathcal{T}} \left(\mathbf{e}_i \mathbf{e}_j^{\intercal} \right) \right\|_F^2} \tag{13a}$$

This measurement gives an element-wise average of non-coverage by $\mathcal{Q}$ in $\mathcal{T}$ which has a maximum value of 1. Second, we measure how much of the subgradient of the nuclear norm at $\mathbf{M}$ is not contained in $\mathcal{Q}$:

$$\nu_{\mathcal{Q}^\perp} := \frac{1}{r} \left\| \mathcal{P}_{\mathcal{Q}^\perp} \left(\mathbf{E} \right) \right\|_F^2 . \tag{14a}$$

where $\mathbf{E} := \sum_{k=1}^{r} \mathbf{u}_k \mathbf{v}_k^\intercal$ and it has a maximum value of 1.

Notice that if $\mathcal{Q}$ covers the entire space of $\mathcal{T}$ (i.e. $\mathcal{Q} \supseteq \mathcal{T}$), then $\mu_{\mathcal{Q}^\perp} = \nu_{\mathcal{Q}^\perp} = 0$ since $\mathbf{E} \in \mathcal{T}$. This extreme case will be important in explaining the significance of our main result, Theorem 1, in regards to how much fewer samples are needed for exact completion. However, this does not mean that no observations are needed because the useful information described above only refers to the information in the r rank-1 matrices but does not say anything about the singular values themselves that need to be determined.

Finally, using the above definitions and assumptions, we can state our theorem on sample complexity with a high-probability matrix completion guarantee.

Theorem 1. *Let $\mathbf{M}$ be an $n_1 \times n_2$ matrix with $n_1 \geq n_2$ such that the following h linear equality constraints are satisfied: $\langle \mathbf{A}^{(l)}, \mathbf{M} \rangle = b^{(l)}$ for all $l \in \{1, \ldots, h\}$. Also, let $\mathbf{M}$ be of rank r and have the following singular value decomposition $\sum_{k=1}^{r} \sigma_k \mathbf{u}_k \mathbf{v}_k^\intercal$ that satisfies Assumptions 1 and 2. Suppose that m entries of $\mathbf{M}$ are sampled uniformly at random. Then there exists a function $F(n_1, n_2, r, \mu_0, \nu_0, \beta, \mu_{\mathcal{Q}^\perp}, \nu_{\mathcal{Q}^\perp}) < \infty$ such that if $m \geq \max\{F, 2\beta n_1 \log n_1\}$ for some $\beta \geq 1$, then the solution to Problem* (11) *is unique and equal to $\mathbf{M}$ with probability at least $1 - 6n_1^{-\beta}$. Specifically, if $\mu_{\mathcal{Q}^\perp} = \nu_{\mathcal{Q}^\perp} = 0$, then there exists constants (C_1, C_2, C_3) for which*

$$F = \max\left\{ C_1 \nu_0^2, C_2 \sqrt{\mu_0 n_2}, C_3 \mu_0 \right\} \beta r n_1 \log n_1 - n_1 n_2 \tag{15a}$$

or if $\mu_{\mathcal{Q}^\perp} = \nu_{\mathcal{Q}^\perp} = 1$, then

$$F = \max\left\{ C_1 \nu_0^2, C_2 \sqrt{\mu_0 n_2}, C_3 \mu_0 \right\} \beta r n_1 \log n_1 . \tag{15b}$$

The proof is given in detail in our extended version [31].

Remark 1. *The sample complexity described by* (15b) *is within $O(\max\{\mu_0^{-\frac{1}{2}} n_2^{\frac{1}{4}}, \mu_0^{-\frac{1}{2}} \nu_0\})$ of [25] for the unconstrained problem.*

This theorem shows us that when $\mu_{\mathcal{Q}^\perp} = \nu_{\mathcal{Q}^\perp} = 0$, i.e. $\mathcal{Q}$ completely covers $\mathcal{T}$, then the reduction in sample complexity is on the order of the size of the matrix $n_1 n_2$, and when $\mu_{\mathcal{Q}^\perp} = \nu_{\mathcal{Q}^\perp} = 1$, i.e. $\mathcal{Q}$ does not cover any of $\mathcal{T}$, there is no reduction. Therefore, this is a preliminary indication that $\mu_{\mathcal{Q}\perp}$ and $\nu_{\mathcal{Q}\perp}$ are useful metrics to characterize the sample complexity reduction from constraints. Precise results for the intermediate cases are subject of future work.

When connecting this theoretical result back to the state estimation problem for a distribution network, it shows us that states which result in a $\mathcal{T}$ that lies in the constraint vector space $\mathcal{Q}$ require significantly less samples for estimation than those that are not.

IV. Performance Evaluation

Using a power system emulator, our goal is to show how incorporating equality constraints based on the physics of the system can improve the accuracy for state estimation.

A. Setup

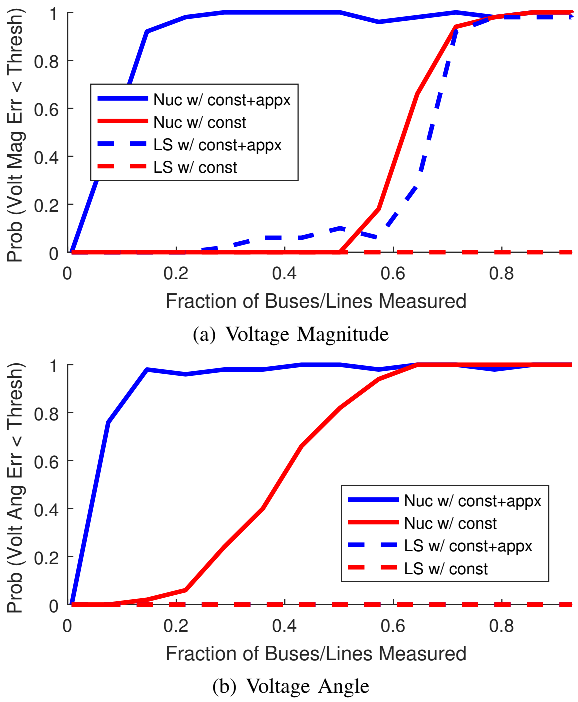
The distribution network data was created using MATPOWER [32] on a 141 bus radial distribution network test case [4]. The underlying matrix $\mathbf{M}$ that represents the state of the power system was formed according to the structure described in Section II-B. Therefore, the state matrix $\mathbf{M}$ has 281 rows and 17 columns. The set of $4(n_\text{b} + n_\text{l}) = 1124$ linear equality constraints (11c) were formed according to the following linear power system equations: (1)-(4), labeled "w/ const" in the figures. An additional set of $n_\text{l} = 140$ linear equality constraints were formed according to the linear approximation equations (9), labeled "w/ const+appx".

To sample the values of the state matrix $\mathbf{M}$, we set that Bus 1 and Bus 80 each have a PMU which can measure all 8 bus state values. The remaining 139 buses and 140 lines were chosen uniformly at random to have standard electrical measurement equipment (e.g. smart meter) that can only read a subset of the state values. When a bus was chosen, only the following 4 values were revealed: the real and reactive power injections, the magnitude of the voltage, and the magnitude of the current injection. When a line was chosen, only the following 5 values were revealed: real and reactive power injections into the line for both the From and To sides of the line, and the magnitude of the current flowing through the line. The uniform sampling was done by taking a random permutation of all the buses and lines together and using the first m buses/lines as the observed samples. For each sample size, 50 different random permutations of the buses and lines were used to do the uniform sampling.

We also solved the widely standard Least Squares (LS) [33] problem as a benchmark by replacing the Nuclear Norm with the Frobenius Norm in Problem (11). To measure the estimation error, the Root Mean Squared Error (RMSE) was taken for voltage magnitude and voltage angle for the unmeasured values. Because all other state quantities can be derived from the complex bus voltages and the physical properties of the power system equipment, our focus in these simulations is on the accuracy of the estimated complex voltages at the buses. The estimated voltage angle is calculated by translating $\text{Re}(V_s)$ and $\text{Im}(V_s)$ from the estimated state matrix into polar form. The estimated voltage magnitude $|V_s|$ is taken directly from the estimated state matrix. The error is calculated by subtracting the estimation from the true value and are only of the unobserved matrix elements.

B. Results

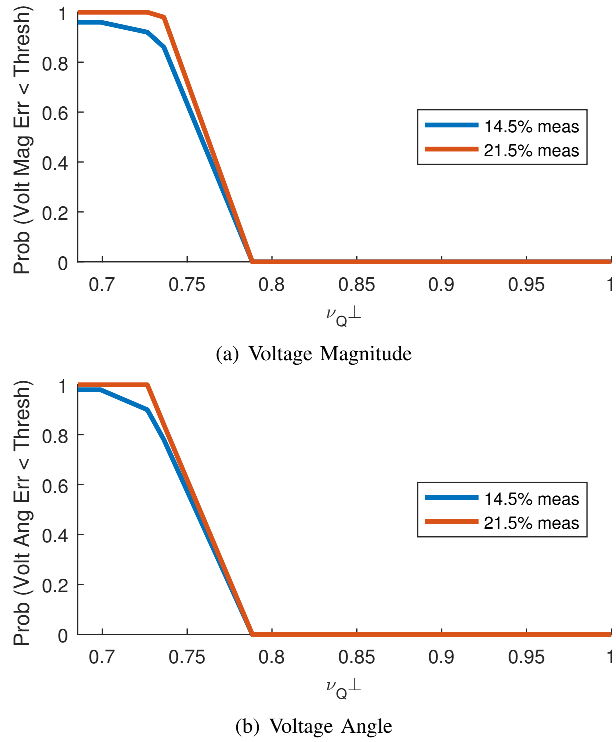
To see how the sample size affects the accuracy of the estimated voltages, we set RMSE thresholds and then counted the fraction of trials tested for each sample size that had RMSEs lower than the threshold. Figure 2 shows the results for error thresholds of 1×10^{-4} pu and 5×10^{-5} degrees for voltage magnitude and voltage angle, respectively. From



(a) Voltage Magnitude

(b) Voltage Angle

Fig. 2. Probability of the estimated Voltage (a) Magnitude RMSE being below a threshold of 1×10^{-4} pu and (b) Angle RMSE being below a threshold of 5×10^{-5} degrees vs. fraction of buses and lines with observed data. For Voltage Angle (b), neither constraint set using the Least Squares method had any trials have error below the threshold.



(a) Voltage Magnitude

(b) Voltage Angle

Fig. 3. Probability of the estimated Voltage (a) Magnitude RMSE being below a threshold of 1×10^{-4} pu and (b) Angle RMSE being below a threshold of 5×10^{-5} degrees vs. $\nu_{Q^\perp}$.

these plots, we can make two strong obesrvations. The first is that the Nuclear Norm method almost always has a higher probability of being more accurate than the Least Squares method for all sample sizes. The second is that the linear approximation equations (9) greatly improve the accuracy of the Nuclear Norm Minimization method to the point that even with only 20% of the buses and lines measured, the unmeasured voltages have over a 90% probability of having their average error be below 1×10^{-4} pu and 5×10^{-5} degrees.

To see how the value of $\nu_{Q^\perp}$ affects the state-estimation, we randomly deleted constraints from the "const+appx" set and solved Problem (11) while measuring $\nu_{Q^\perp}$. Figure 3 shows the results for the same error thresholds as before with two different sample sizes of 14.5% and 21.5% of the buses and lines. As constraints are added, the value of $\nu_{Q^\perp}$ decreases. We can observe a threshold value of $\nu_{Q^\perp}$ at 0.79 before the added constraints help to increase the probability have having small error. This gives evidence to the idea that the constraint set must achieve a small enough $\nu_{Q^\perp}$ before it can be fully utilized with a small sample size.

V. Related Work

In traditional state estimation, the focus is mainly on transmission networks that have an abundance of measurement equipment so that the focus is on how to remove bad data using weighted least-squares techniques [33]. For distribution networks that are measurement poor, these techniques cannot be used since they require full observability. Matrix Completion has only been recently considered for power system state estimation with the use of PMUs [34]. The structure of our problem mainly follows that of [24] to take advantage of power system physics to add information. While we focus our analysis on a more theoretical perspective of sample complexity, [24] uses a detailed simulations to measure estimation errors under different low-observablility scenarios. The work of [34], [35], [36], [37] focuses on the time correlation of a single state variable type by using one of the matrix dimensions to represent time, as compared to the state variable type in our problem and [24] to the focus on the correlation between state variable types at a single moment.

The low-rank matrix completion theoretical framework from [25] used the Bernoulli sampling model to bound the failure probability for uniform sampling. While we also used this sampling model, however their framework wasn't flexible enough or didn't have the intended purpose to include constraints. Linear equality constraints were used to convey information for matrix completion in [38] instead of sampling. However, compared to our problem, theirs modeled the constraints themselves as random instead of as a permanent feature of the matrix being completed.

VI. Conclusion

In this paper, we develop a method for distribution network state estimation which has the characteristic of being *underdetermined* as opposed to the traditional overdetermined state estimation problem found in transmission networks. Our method generalizes standard low-rank matrix completion techniques so that the physical properties governing a power system can reduce the number of samples needed to estimate the state, which was unable to be done the existing methods. The sample complexity for high-probability exact matrix completion was proved for the constrained matrix completion problem using nuclear norm minimization. This shows how the additional information obtained from linear equality constraints can reduce the number of samples needed to exactly recover the underlying matrix. The method was tested with real-world data on a 141 bus distribution network test case and shows that the estimation error for voltage magnitude and angle at each bus can be significantly reduced. Without the reduction, we wouldn't be able to accurately recover the matrix and thus also wouldn't be able to carefully control the power system.

VII. Acknowledgements

This work was authored in part by the National Renewable Energy Laboratory, managed and operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36- 08GO28308. This work was supported in part by the Laboratory Directed Research and Development (LDRD) Program at NREL. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes. This work was also partially funded by NSF grants CNS-1464388, CNS-1617698, CNS-1730128, CNS-1717558, DGE-1633299.

References

[1] M. E. Baran and F. F. Wu, "Network reconfiguration in distribution systems for loss reduction and load balancing," *IEEE Transactions on Power delivery*, vol. 4, no. 2, pp. 1401–1407, 1989.

[2] D. Das, "Optimal placement of capacitors in radial distribution system using a fuzzy-ga method," *International Journal of Electrical Power & Energy Systems*, vol. 30, no. 6-7, pp. 361–367, 2008.

[3] D. Das, D. Kothari, and A. Kalam, "Simple and efficient method for load flow solution of radial distribution networks," *International Journal of Electrical Power & Energy Systems*, vol. 17, no. 5, pp. 335–346, 1995.

[4] H. Khodr, F. Olsina, P. De Oliveira-De Jesus, and J. Yusta, "Maximum savings approach for location and sizing of capacitors in distribution systems," *Electric Power Systems Research*, vol. 78, no. 7, pp. 1192–1203, 2008.

[5] T. D. Liacco, "The role of state estimation in power system operation," *IFAC Proceedings Volumes*, vol. 15, no. 4, pp. 1531–1533, 1982.

[6] A. Abur and A. Gomez-Exposito, *Power System State Estimation: Theory and Implementation*. CRC press, 2004.

[7] A. Meliopoulos, B. Fardanesh, and S. Zelingher, "Power system state estimation: modeling error effects and impact on system operation," in *hicss*. IEEE, 2001, p. 2012.

[8] K. Dehghanpour, Z. Wang, J. Wang, Y. Yuan, and F. Bu, "A survey on state estimation techniques and challenges in smart distribution systems," *IEEE Transactions on Smart Grid*, 2018.

[9] J. Sexauer, P. Javanbakht, and S. Mohagheghi, "Phasor measurement units for the distribution grid: Necessity and benefits," in *Innovative Smart Grid Technologies (ISGT), 2013 IEEE PES*. IEEE, 2013, pp. 1–6.

[10] S. Bolognani, R. Carli, and M. Todescato, "State estimation in power distribution networks with poorly synchronized measurements," in *Decision and Control (CDC), 2014 IEEE 53rd Annual Conference on*. IEEE, 2014, pp. 2579–2584.

[11] M. E. Baran, "Challenges in state estimation on distribution systems," in *2001 Power Engineering Society Summer Meeting. Conference Proceedings (Cat. No.01CH37262)*, vol. 1, July 2001, pp. 429–433 vol.1.

[12] J. Driesen and F. Katiraei, "Design for distributed energy resources," *IEEE Power and Energy Magazine*, vol. 6, no. 3, 2008.

[13] A. Kulmala, S. Repo, and P. Järventausta, "Coordinated voltage control in distribution networks including several distributed energy resources," *IEEE Transactions on Smart Grid*, vol. 5, no. 4, pp. 2010–2020, 2014.

[14] R. Syahputra, I. Robandi, and M. Ashari, "Optimal distribution network reconfiguration with penetration of distributed energy resources," in *Information Technology, Computer and Electrical Engineering (ICITACEE), 2014 1st International Conference on*. IEEE, 2014, pp. 388–393.

[15] I.-K. Song, W.-W. Jung, J.-Y. Kim, S.-Y. Yun, J.-H. Choi, and S.-J. Ahn, "Operation schemes of smart distribution networks with distributed energy resources for loss reduction and service restoration," *IEEE Transactions on Smart Grid*, vol. 4, no. 1, pp. 367–374, 2013.

[16] F. Katiraei and J. R. Aguero, "Solar pv integration challenges," *IEEE Power and Energy Magazine*, vol. 9, no. 3, pp. 62–71, 2011.

[17] R. H. Inman, H. T. Pedro, and C. F. Coimbra, "Solar forecasting methods for renewable energy integration," *Progress in energy and combustion science*, vol. 39, no. 6, pp. 535–576, 2013.

[18] C. A. Hill, M. C. Such, D. Chen, J. Gonzalez, and W. M. Grady, "Battery energy storage for enabling integration of distributed solar power generation," *IEEE Transactions on smart grid*, vol. 3, no. 2, pp. 850–857, 2012.

[19] G. Energy, "Western wind and solar integration study," Citeseer, Tech. Rep., 2010.

[20] D. Della Giustina, M. Pau, P. A. Pegoraro, F. Ponci, and S. Sulis, "Electrical distribution system state estimation: measurement issues and challenges," *IEEE Instrumentation & Measurement Magazine*, vol. 17, no. 6, pp. 36–42, 2014.

[21] Y.-F. Huang, S. Werner, J. Huang, N. Kashyap, and V. Gupta, "State estimation in electric power grids: Meeting new challenges presented by the requirements of the future grid," *IEEE Signal Processing Magazine*, vol. 29, no. 5, pp. 33–43, 2012.

[22] A. Alarcon-Rodriguez, G. Ault, and S. Galloway, "Multi-objective planning of distributed energy resources: A review of the state-of-the-art," *Renewable and Sustainable Energy Reviews*, vol. 14, no. 5, pp. 1353–1366, 2010.

[23] P. Basak, S. Chowdhury, S. H. nee Dey, and S. Chowdhury, "A literature review on integration of distributed energy resources in the perspective of control, protection and stability of microgrid," *Renewable and Sustainable Energy Reviews*, vol. 16, no. 8, pp. 5545–5556, 2012.

[24] P. L. Donti, Y. Liu, A. J. Schmitt, A. Bernstein, and Y. Zhang, "Matrix completion for low-observability voltage estimation," *arXiv preprint arXiv:1801.09799*, 2019.

[25] E. J. Candès and B. Recht, "Exact matrix completion via convex optimization," *Foundations of Computational mathematics*, vol. 9, no. 6, p. 717, 2009.

[26] M. Simchowitz, H. Mania, S. Tu, M. I. Jordan, and B. Recht, "Learning without mixing: Towards a sharp analysis of linear system identification," in *Proceedings of the 31st Conference On Learning Theory*, ser. Proceedings of Machine Learning Research, vol. 75. PMLR, 06–09 Jul 2018, pp. 439–473.

[27] A. Rantzer, "Concentration bounds for single parameter adaptive control," in *2018 Annual American Control Conference (ACC)*, June 2018, pp. 1862–1866.

[28] E. J. Candes and Y. Plan, "Matrix completion with noise," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 925–936, 2010.

[29] M. Baran and F. F. Wu, "Optimal sizing of capacitors placed on a radial distribution system," *IEEE Transactions on power Delivery*, vol. 4, no. 1, pp. 735–743, 1989.

[30] M. Fazel, "Matrix rank minimization with applications," Ph.D. dissertation, PhD thesis, Stanford University, 2002.

[31] J. Comden, M. Colombino, A. Bernstein, and Z. Liu, "Sample complexity of power system state estimation using matrix completion," *arXiv preprint arXiv:1905.01789*, 2019.

[32] R. D. Zimmerman, C. E. Murillo-Sánchez, R. J. Thomas *et al.*, "Matpower: Steady-state operations, planning, and analysis tools for power

systems research and education," *IEEE Transactions on power systems*, vol. 26, no. 1, pp. 12–19, 2011.

[33] A. Monticelli, "Electric power system state estimation," *Proceedings of the IEEE*, vol. 88, no. 2, pp. 262–282, 2000.

[34] M. Wang, J. H. Chow, P. Gao, X. T. Jiang, Y. Xia, S. G. Ghiocel, B. Fardanesh, G. Stefopolous, Y. Kokai, N. Saito *et al.*, "A low-rank matrix approach for the analysis of large amounts of power system synchrophasor data," in *System Sciences (HICSS), 2015 48th Hawaii International Conference on*. IEEE, 2015, pp. 2637–2644.

[35] P. Gao, M. Wang, S. G. Ghiocel, J. H. Chow, B. Fardanesh, and G. Stefopoulos, "Missing data recovery by exploiting low-dimensionality in power system synchrophasor measurements," *IEEE Transactions on Power Systems*, vol. 31, no. 2, pp. 1006–1013, 2016.

[36] C. Genes, I. Esnaola, S. M. Perlaza, L. F. Ochoa, and D. Coca, "Recovering missing data via matrix completion in electricity distribution systems," in *Signal Processing Advances in Wireless Communications (SPAWC), 2016 IEEE 17th International Workshop on*. IEEE, 2016, pp. 1–6.

[37] ——, "Robust recovery of missing data in electricity distribution systems," *IEEE Transactions on Smart Grid*, 2018.

[38] B. Recht, M. Fazel, and P. A. Parrilo, "Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization," *SIAM review*, vol. 52, no. 3, pp. 471–501, 2010.