



Matrix Completion Using Alternating Minimization for Distribution System State Estimation

Preprint

Yajing Liu,¹ April Sagan,² Andrey Bernstein,¹ Rui Yang,¹ Xinyang Zhou,¹ and Yingchen Zhang¹

¹ *National Renewable Energy Laboratory*

² *Rensselaer Polytechnic Institute*

To be presented at the IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (IEEE SmartGridComm) November 11–13, 2020

**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-5D00-74961
October 2020



Matrix Completion Using Alternating Minimization for Distribution System State Estimation

Preprint

Yajing Liu,¹ April Sagan,² Andrey Bernstein,¹ Rui Yang,¹ Xinyang Zhou,¹ and Yingchen Zhang¹

¹ *National Renewable Energy Laboratory*

² *Rensselaer Polytechnic Institute*

Suggested Citation

Liu, Yajing, April Sagan, Andrey Bernstein, Rui Yang, Xinyang Zhou, and Yingchen Zhang. 2020. *Matrix Completion Using Alternating Minimization for Distribution System State Estimation: Preprint*. Golden, CO: National Renewable Energy Laboratory. NREL/CP-5D00-74961. <https://www.nrel.gov/docs/fy21osti/74961.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-5D00-74961
October 2020

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. Funding provided by U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Solar Energy Technologies Office. The views expressed herein do not necessarily represent the views of the DOE or the U.S. Government.

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

U.S. Department of Energy (DOE) reports produced after 1991 and a growing number of pre-1991 documents are available free via www.OSTI.gov.

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.

Matrix Completion Using Alternating Minimization for Distribution System State Estimation

Yajing Liu*, April Sagan†, Andrey Bernstein*, Rui Yang*, Xinyang Zhou*, Yingchen Zhang*

*National Renewable Energy Laboratory, Golden, USA

†Rensselaer Polytechnic Institute, Troy, USA

Emails: {yajing.liu, andrey.bernstein, rui.yang, xinyang.zhou, yingchen.zhang}@nrel.gov, sagana@rpi.edu

Abstract—This paper examines the problem of state estimation in power distribution systems under low-observability conditions. The recently proposed constrained matrix completion method which combines the standard matrix completion method and power flow constraints has been shown to be effective in estimating voltage phasors under low-observability conditions using single-snapshot information. However, the method requires solving a semidefinite programming (SDP) problem, which becomes computationally infeasible for large systems and if multiple-snapshot (time-series) information is used. This paper proposes an efficient algorithm to solve the constrained matrix completion problem with time-series data. This algorithm is based on reformulating the matrix completion problem as a bilinear (non-convex) optimization problem, and applying the alternating minimization algorithm to solve this problem. This paper proves the summable convergence of the proposed algorithm, and demonstrates its efficacy and scalability via IEEE 123-bus system and a real utility feeder system. This paper also explores the value of adding more data from the history in terms of computation time and estimation accuracy.

I. INTRODUCTION

State estimation in power systems is critical for maintaining normal and secure operating conditions. In power transmission systems, the weighted least squares (WLS) method [1] is well developed and widely used by utilities for state estimation; however, the limited availability of measurements renders the WLS inapplicable in power distribution systems. The requirement for state estimation in distribution systems is becoming stringent because the increased penetration of distributed energy resources—such as solar photovoltaics, wind turbines, and energy storage systems—on distribution systems introduces bidirectional power flows that might impact the system responses to various types of disturbances.

Much work has been done in an attempt to address the low observability state estimation problem. Reference [2]

Financial support from the National Science Foundation under Grant Number DMS-1736326 is gratefully acknowledged. 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. Funding provided by the U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Solar Energy Technologies Office. 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 non-exclusive, 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.

determined the optimal minimum measurement locations and used the difference between the measured and calculated voltages and complex powers to obtain the voltage profile of the whole network. Reference [3] used the voltage sensitivity at each bus to the load at all buses to determine the voltage given the measured load at each location. References [4]–[7] applied neural networks, which do not require system models, to estimate the voltage in the distribution grid; however, these approaches still require the installation of significant numbers of phasor measurement units (PMUs) when considering large systems. References [4], [8]–[10] used pseudo-measurements, which are estimates based on historical data to overcome the requirement of the installation of more meters; however, it is known that the pseudo-measurements typically have larger measurement errors than the metered measurements.

An alternative approach that does not require installing new sensors or explicit computation of pseudo-measurements was recently proposed by [11]. This method leverages the standard matrix completion method [12] and augments it with linearized power flow constraints to acknowledge for the physical network constraints. We call this method *constrained matrix completion method*. This method was shown to be effective in estimating voltage phasors using single-snapshot information. However, this method requires solving a semidefinite programming (SDP) problem, which becomes computationally infeasible for large systems or if multiple-snapshot (time-series) information is used.

In fact, temporal correlation (dependency between measurements at different time steps) exists among data in addition to the spatial correlation (dependency between measurements at different locations) and the correlation among measurement types (characterized by power flow equations). By modeling both the temporal and spatial correlation of the data, [13], [14] apply the standard matrix completion method to recover missing PMU data, and [15] applies the standard matrix completion method with Bayesian estimation to recover missing low-voltage distribution system data.

This paper proposes an efficient algorithm to solve the constrained matrix completion problem with time-series data. To this end, this paper formulates matrix X using time-series data by considering both the spatial and temporal correlation of the data. Similar to [11], this paper leverages a linear approximation of the power flow equations as constraints

to characterize the correlation among different measurement types. Then, this paper applies the alternating minimization method, which was a winner in the Netflix Challenge [16], to solve the proposed matrix completion problem for state estimation in multiphase distribution systems.

To apply the alternating minimization method, matrix X is written in a bilinear form, i.e., $X = UV$, and then the algorithm finds the best U and V in an alternating fashion. Because of the bilinear term in the objective function, the problem becomes nonconvex with respect to any variable and thus the alternating minimization method is not guaranteed to converge to the global optima. However, results on the convergence of the method were established under some regularity conditions in [17], [18]. This paper proves that the objective function satisfies these conditions, which guarantee the convergence of the proposed method.

The rest of this paper is organized as follows. Section II introduces the standard matrix completion model, the formulation of the data matrix using time-series data, the linearized power flow constraints, and the full matrix completion model for state estimation in multiphase distribution systems. The alternating minimization algorithm and the summable convergence are presented in Section III. Simulation results on the IEEE 123-bus system and a real utility feeder are shown in Section IV. Section V concludes the paper.

II. PROBLEM FORMULATION

This section first reviews the standard matrix completion formulation, and then introduces how to build the data matrix using the time-series data in multiphase distribution systems. Next, the linearized power flow constraints are introduced. Finally, the full matrix completion model for state estimation in multiphase distribution systems is derived, which combines the standard matrix completion model with the linear power flow constraints.

A. Standard Matrix Completion

Let $M \in \mathbb{R}^{m \times n}$ be the matrix we wish to reconstruct, and let Ω denote the set of indices for known elements of M . Define the observation operator $P_\Omega : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{m \times n}$ as: $P_\Omega(M)_{ij} = M_{ij}$ for $(i, j) \in \Omega$; 0 otherwise.

The goal is to recover M from $P_\Omega(M)$, under the assumption that M is low rank. Because the rank function is nonconvex and the tightest convex relaxation of the rank function is the nuclear norm function [12], we consider the following regularized matrix completion model:

$$\min_{X \in \mathbb{R}^{m \times n}} \|X\|_* + \frac{\mu}{2} \|P_\Omega(X) - P_\Omega(M)\|_F^2, \quad (1)$$

where X is the decision variable, $\|X\|_* = \sum_{i=1}^{\min\{m,n\}} \sigma_i(X)$ is the nuclear norm of X with $\sigma_i(X)$ denoting the i th largest singular value of X , and $\mu > 0$ is a weight parameter.

B. Matrix Formulation in Multiphase Distribution Systems

For the scope of this paper, we consider a general three-phase distribution network consisting of one slack bus and a given number of multiphase PQ buses.

Matrix completion was shown effective for state estimation in general three-phase distribution systems by [11], which formulated the data matrix M using single-snapshot information. Considering the temporal correlation of the data, we set up the data matrix M using time-series data.

Assume that the voltage phasor and other measurements at the slack bus are known. So we use only measurements at nonslack buses to form the data matrix. Let \mathcal{P} denote the set of phases at all nonslack buses and $|\mathcal{P}|$ be the number of all phases at nonslack buses. The measurements we use in this matrix are the real and imaginary parts of voltage phasor, voltage magnitude, and the active and reactive power injection at each phase of nonslack buses. Consider a time series $t = 1, \dots, T$. Let M^t denote the measurement matrix at time t such that each column represents a phase and each row represents a quantity relevant to the phase. To be specific, for each phase $i \in \mathcal{P}$, the corresponding column of M^t is of the form:

$$[\Re(v_i), \Im(v_i), |v_i|, \Re(s_i), \Im(s_i)]^\top,$$

where $\Re(\cdot)$ and $\Im(\cdot)$ are the real part and imaginary part of a complex variable, respectively, \top is the transpose notation, $v = [v_1, \dots, v_{|\mathcal{P}|}]^\top \in \mathbb{C}^{|\mathcal{P}|}$ is the vector containing voltage phasors at each phase of nonslack buses, and $s = [s_1, \dots, s_{|\mathcal{P}|}]^\top \in \mathbb{C}^{|\mathcal{P}|}$ is the vector of power injections at each phase of nonslack buses. Then the matrix M is constructed by

$$M = [M^1; M^2; \dots; M^T] \in \mathbb{R}^{m \times n}, \quad (2)$$

where T is the length of the time series, and $m = 5T, n = |\mathcal{P}|$. Note that the data matrix is not limited to the structure proposed above; it can accommodate any available measurements if these measurements are correlated such that M has the (approximate) low-rank property.

Because the rows and/or columns of the matrix M are related by temporal correlation, spatial correlation, or power flow equations, such a matrix has low rank. Fig. 1 shows that the low-rank property holds for the IEEE 123-bus system with one-time step, two-time step, and three-time step data formulation.

C. Linear Power Flow Constraints

As in [11], we use the linear approximation of power flows as constraints to characterize the correlation among different measurement types, which provides more information about the system, and thus likely requires fewer measurements to recover the matrix than the standard matrix completion formulation.

Recall that we use \mathcal{P} , $v \in \mathbb{C}^{|\mathcal{P}|}$, and $s \in \mathbb{C}^{|\mathcal{P}|}$ to denote the set of phases, the vector of voltage phasors, and the vector of power injections at all nonslack buses, respectively. We employ approximations of the form:

$$v \approx B \begin{bmatrix} \Re(s) \\ \Im(s) \end{bmatrix} + w, \quad (3a)$$

$$|v| \approx C \begin{bmatrix} \Re(s) \\ \Im(s) \end{bmatrix} + |w|, \quad (3b)$$

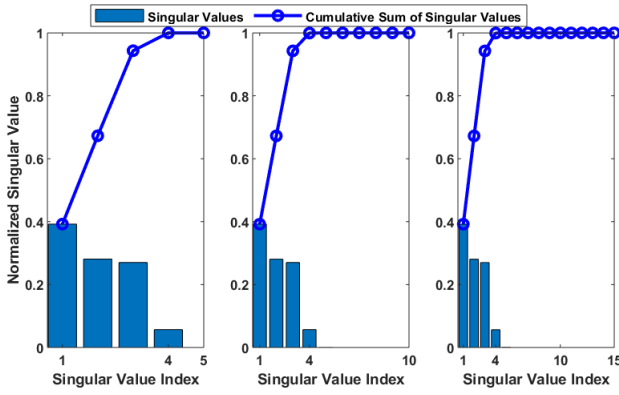


Fig. 1: Singular values of data matrices for the IEEE 123-bus system with one-time step, two-time step, and three-time step data matrix formulation. The data matrices are of size 5×260 , 10×260 , and 15×260 , respectively. The bars show the individual singular values (normalized by the singular value sum), and the circles show the cumulative sums of (normalized) singular values. The first four largest singular values comprise 99% of the singular value sum.

where the coefficients $B, C \in \mathbb{C}^{|\mathcal{P}| \times 2|\mathcal{P}|}$, $w \in \mathbb{C}^{|\mathcal{P}|}$ are derived by the power flow linearization methods developed in, e.g., [19]–[21]. Write (3) in the following equivalent form with $A_1, A_2, A_3, A_4, C_1, C_2 \in \mathbb{R}^{|\mathcal{P}| \times |\mathcal{P}|}$:

$$\Re(v) \approx [A_1 \ A_2] \begin{bmatrix} \Re(s) \\ \Im(s) \end{bmatrix} + \Re(w), \quad (4a)$$

$$\Im(v) \approx [A_3 \ A_4] \begin{bmatrix} \Re(s) \\ \Im(s) \end{bmatrix} + \Im(w), \quad (4b)$$

$$|v| \approx [C_1 \ C_2] \begin{bmatrix} \Re(s) \\ \Im(s) \end{bmatrix} + |w|, \quad (4c)$$

where $[A_1 \ A_2] = \Re(B)$, $[A_3 \ A_4] = \Im(B)$, and $[C_1 \ C_2] = C$.

We use v^t, s^t to denote the corresponding v, s in (4) at time slot t . Assume that v_0 (the slack bus voltage phasor) and the system topology are the same at different time steps, and then by using the linearization method in [19], $A_1, A_2, A_3, A_4, C_1, C_2$ are the same at different time steps. Hence, we have the linear approximation of voltage phasor and magnitude at time t in (4) with v, s being replaced by v^t and s^t .

For simplicity of expression in the sequel, we use the following model to express this linear model at time $t = 1, \dots, T$:

$$y \approx Ax + b, \quad (5)$$

where:

$$y = \begin{bmatrix} \Re(v^1) \\ \Im(v^1) \\ |v^1| \\ \vdots \\ \Re(v^T) \\ \Im(v^T) \\ |v^T| \end{bmatrix}, \quad A = \begin{bmatrix} A_1 & A_2 & 0 & \cdots & 0 & 0 \\ A_3 & A_4 & 0 & \cdots & 0 & 0 \\ C_1 & C_2 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 & 0 \\ 0 & 0 & 0 & \cdots & A_1 & A_2 \\ 0 & 0 & 0 & \cdots & A_3 & A_4 \\ 0 & 0 & 0 & \cdots & C_1 & C_2 \end{bmatrix},$$

$$x = [\Re(s^1)^\top \ \Im(s^1)^\top \ \dots \ \Re(s^T)^\top \ \Im(s^T)^\top]^\top,$$

and

$$b = [\Re(w)^\top \ \Im(w)^\top \ |w|^\top \ \dots \ \Re(w)^\top \ \Im(w)^\top \ |w|^\top]^\top$$

with $A \in \mathbb{R}^{3T|\mathcal{P}| \times 2T|\mathcal{P}|}$, $x \in \mathbb{R}^{2T|\mathcal{P}|}$, and $y, b \in \mathbb{R}^{3T|\mathcal{P}|}$.

The linear model (5) will be incorporated into the matrix completion formulation in the following subsection.

D. The Regularized Matrix Completion for State Estimation

By incorporating the linear power flow model (5) into problem (1) as a regularization term, we obtain the following optimization problem:

$$\begin{aligned} \min \quad & \|X\|_* + \frac{\mu}{2} \|P_\Omega(X - M)\|_F^2 + \frac{\nu}{2} \|y - (Ax + b)\|_2^2 \\ \text{over } & X \in \mathbb{R}^{m \times n}, y \in \mathbb{R}^{\frac{3}{2}mn}, x \in \mathbb{R}^{\frac{2}{2}mn}, \\ \text{s.t. } & y = [a_1^\top X \ a_2^\top X \ a_3^\top X \ \dots \ a_{3T-2}^\top X \ a_{3T-1}^\top X \ a_{3T}^\top X]^\top, \\ & x = [c_1^\top X \ c_2^\top X \ \dots \ c_{2T-1}^\top X \ c_{2T}^\top X]^\top, \end{aligned} \quad (6)$$

where, $\nu > 0$ is a weight parameter, $m = 5T, n = |\mathcal{P}|$, $a_{3(t-1)+i} = e_{5(t-1)+i}$ and $c_{2(t-1)+i} = e_{5(t-1)+3+i}$ are standard basis vectors in \mathbb{R}^m , in which t denotes the time step number and i denotes the i th a or c at the t -th time step. For example, if $a_1, t = 1, i = 1$, then, $a_1 = e_1 \in \mathbb{R}^m$; if $c_1, t = 1, i = 1$, then $c_1 = e_4 \in \mathbb{R}^m$.

III. ALTERNATING MINIMIZATION FOR SOLVING THE FULL MATRIX COMPLETION MODEL

This section first proposes the alternating minimization algorithm for the full matrix completion model in multiphase distribution system state estimation, and then proves the summable convergence of the proposed algorithm.

A. Alternating Minimization Algorithm

Any matrix $X \in \mathbb{R}^{m \times n}$ of a rank up to r has a matrix product $X = UV$ form with $U \in \mathbb{R}^{m \times r}$ and $V \in \mathbb{R}^{r \times n}$, and its nuclear norm can be expressed by the Frobenius norm of U and V as follows [22]:

$$\|X\|_* := \arg \min_{U \in \mathbb{R}^{m \times r}, V \in \mathbb{R}^{r \times n}} \frac{1}{2} (\|U\|_F^2 + \|V\|_F^2) \quad (7)$$

s. t. $X = UV.$

Note that for sufficiently small r ($r \ll \min\{m, n\}$), using this characterization of the nuclear norm allows us to dramatically reduce the size of the problem. We now use (7) to reformulate (6) as follows:

$$\begin{aligned} \min_{\substack{U \in \mathbb{R}^{m \times r}, \\ V \in \mathbb{R}^{r \times n}}} \quad & \frac{1}{2} (\|U\|_F^2 + \|V\|_F^2) + \frac{\mu}{2} \|P_\Omega(UV - M)\|_F^2 \\ & + \frac{\nu}{2} \|f_1(UV) - (Af_2(UV) + b)\|_2^2, \\ \text{s.t. } & f_1(UV) = [a_1^\top X \ a_2^\top X \ a_3^\top X \ \dots \ a_{3T-2}^\top X \ a_{3T-1}^\top X \ a_{3T}^\top X]^\top, \\ & f_2(UV) = [c_1^\top X \ c_2^\top X \ \dots \ c_{2T-1}^\top X \ c_{2T}^\top X]^\top, \\ & X = UV. \end{aligned} \quad (8)$$

The alternating minimization algorithm that updates the variables in an alternating fashion is proved to be one of the most accurate and effective methods for solving the matrix completion problem [23]. For our matrix completion

formulation in state estimation, the alternating minimization algorithm solves for U and V alternatively while fixing the other factor. The pseudo-code of the alternating minimization algorithm for the formulation (8) is given in Algorithm 1.

Algorithm 1 Alternating Minimization Algorithm for Matrix Completion in Distribution System State Estimation

Input: Set of known indices Ω , measurement matrix M , linear model A, b , and the number of iterations N .

Initialize: Calculate the rank r singular value decomposition $[\tilde{U}\tilde{\Sigma}\tilde{V}] = \text{SVD}_r(P_\Omega(M))$, and set $V^{(0)} = \tilde{\Sigma}^{\frac{1}{2}}\tilde{V}$.

1: **for** $k = 1, \dots, N$ **do**

2:

$$U^{(k)} = \arg \min_{U \in \mathbb{R}^{m \times r}} \|U\|_F^2 + \mu \|P_\Omega(UV^{(k-1)} - M)\|_F^2 + \nu \|f_1(UV^{(k-1)}) - (Af_2(UV^{(k-1)}) + b)\|_2^2$$

3:

$$V^{(k)} = \arg \min_{V \in \mathbb{R}^{r \times n}} \|V\|_F^2 + \mu \|P_\Omega(U^{(k)}V - M)\|_F^2 + \nu \|f_1(U^{(k)}V) - (Af_2(U^{(k)}V) + b)\|_2^2$$

4: **end for**

5: **return** $X = U^{(N)}V^{(N)}$

B. Convergence of the Alternating Minimization Algorithm

Because of the bilinear form $X = UV$, problem (8) is not jointly convex with respect to any variable. In general, there is no guarantee that a stationary point of a nonconvex problem will coincide with the global optimum. However, [18] proved the summable convergence of the alternating minimization algorithm if the objective function satisfies the following four conditions: 1) it is continuous in its domain and always bigger than negative infinity; 2) it has a Nash equilibrium; 3) it is strongly convex with respect to each variable; and 4) it satisfies the Kurdyka-Łojasiewicz property (please see Definition 2.5 in [18] for the definition). We show below that our objective function satisfies these four conditions; thus, we have the following proposition.

Proposition 1. *The alternating minimization algorithm (Algorithm 1) satisfies the summable convergence, i.e.: $\sum_{k=1}^{+\infty} \|U^{(k+1)} - U^{(k)}\|_F + \|V^{(k+1)} - V^{(k)}\|_F < +\infty$.*

Proof. The objective function f can be written as follows:

$$f(U, V) = \frac{1}{2} (\|U\|_F^2 + \|V\|_F^2) + \frac{\mu}{2} \|P_\Omega(UV - M)\|_F^2 + \frac{\nu}{2} \|f_1(UV) - (Af_2(UV) + b)\|_2^2, \quad (9)$$

where by letting $X = UV$,

$$\begin{aligned} & \|f_1(UV) - (Af_2(UV) + b)\|_2^2 \\ &= \sum_{i=1}^T \|X^\top a_{3i-2} - (A_1 X^\top c_{2i-1} + A_2 X^\top c_{2i}) - \Re(w)\|_2^2 \\ &+ \sum_{i=1}^T \|X^\top a_{3i-1} - (A_3 X^\top c_{2i-1} + A_4 X^\top c_{2i}) - \Im(w)\|_2^2 \\ &+ \sum_{i=1}^T \|X^\top a_{3i} - (C_1 X^\top c_{2i-1} + C_2 X^\top c_{2i}) - |w|\|_2^2. \quad (10) \end{aligned}$$

First, the domain of the objective function is $(\mathcal{U}, \mathcal{V}) = \{(U, V) : U \in \mathbb{R}^{m \times r}, V \in \mathbb{R}^{r \times n}\}$. By the form of $f(U, V)$ in (9), we have that $f(U, V)$ is continuous in its domain $(\mathcal{U}, \mathcal{V})$, and it is non-negative for any $(U, V) \in (\mathcal{U}, \mathcal{V})$.

Second, to prove that (8) has a Nash point, we prove that (8) has a global minimizer, which by definition is a Nash point. By the definition of $f(U, V)$ in (9), the following expression holds: $\lim_{\|U\| \rightarrow +\infty, \|V\| \rightarrow +\infty} f(U, V) = +\infty$, which means that $f(U, V)$ is coercive [24]. Then, by Theorem 2 in [24], problem (8) has a global minimizer.

Third, we prove that $f(U, V)$ is strongly convex with respect to U and V , respectively. To prove that $f(U, V)$ is strongly convex with respect to U , we prove that $\frac{\partial^2 f(U, V)}{\partial U^2} \succ 0$ as follows.

Let $g(U) := \|f_1(UV) - (Af_2(UV) + b)\|_2^2$ for a fixed V . Then we have

$$\begin{aligned} \frac{\partial^2 f(U, V)}{\partial U^2} &= I_{mr \times mr} + \mu (V \otimes I_{m \times m})^\top R_\Omega (V \otimes I_{m \times m}) \\ &+ \frac{\nu}{2} \frac{\partial^2 g}{\partial U^2}, \quad (11) \end{aligned}$$

where $R_\Omega \in \mathbb{R}^{mr \times mr}$ is the diagonal matrix such that $(R_\Omega)_{ii} = 1$ if $((i \bmod n) + 1, \lceil \frac{i}{m} \rceil) \in \Omega$ and 0 otherwise.

By the form of g in (10), we focus only at the second-order partial derivative of the first term in the first summation term in (10); treatment of the other terms can be derived similarly. Let g_1 denote the first term in the first summation term in (10). We will prove that g_1 is a convex function by definition, then using that $g_1 : \mathbb{R}^{m \times r} \rightarrow \mathbb{R}$ is convex if and only if $\frac{\partial^2 g_1}{\partial U^2} \succeq 0$ to derive that $\frac{\partial^2 g_1}{\partial U^2} \succeq 0$.

By (10), we have that:

$$g_1(U) = \|A_1 V^\top U^\top c_1 + A_2 V^\top U^\top c_2 - V^\top U^\top a_1 + \Re(w)\|_2^2.$$

Because \mathcal{U} denotes the set of all matrices of size $m \times r$, we have that \mathcal{U} is a convex set. Then by the form of $g_1(U)$, it is easy to derive that $\forall U_1, U_2 \in \mathcal{U}, \forall t \in [0, 1]: g_1(tU_1 + (1-t)U_2) \leq tg_1(U_1) + (1-t)g_1(U_2)$.

Hence, g_1 is convex, implying that $\frac{\partial^2 g_1}{\partial U^2} \succeq 0$.

Because all the other terms in (10) share the similar form with g_1 , we can similarly prove that all the other terms in (10) are convex functions with respect to U ; thus, we have that $\frac{\partial^2 g}{\partial U^2} \succeq 0$. By (11), we have that $I_{mr \times mr} \succ 0$, and $\mu (V \otimes I_{m \times m})^\top R_\Omega (V \otimes I_{m \times m}) \succeq 0$; therefore, for $\mu > 0, \nu > 0, \frac{\partial^2 g}{\partial U^2} \succeq 0$, we have $\frac{\partial^2 f(U, V)}{\partial U^2} \succ 0$.

Hence, we conclude that f is strongly convex with respect to U . The proof that f is strongly convex with respect to V is similar and is omitted for brevity.

By [18], any real analytic function satisfies the Kurdyka-Łojasiewicz property. And by definition of the real analytic function, the objective function (9) is real analytic; thus, (9) satisfies the Kurdyka-Łojasiewicz property.

This concludes the proof. \square

IV. SIMULATION RESULTS

This section demonstrates Algorithm 1 on the IEEE 123-bus test system [25] and a real utility feeder system. The two systems are both three-phase unbalanced radial distribution

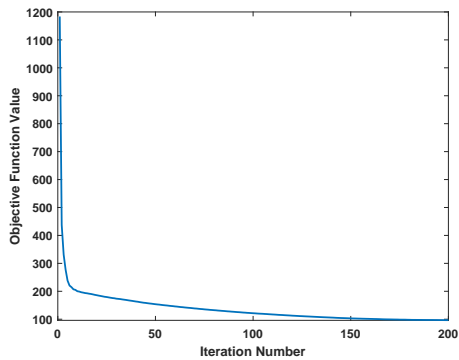


Fig. 2: Convergence of the alternating minimization algorithm with one-time step data matrix formulation and 50% available measurements.

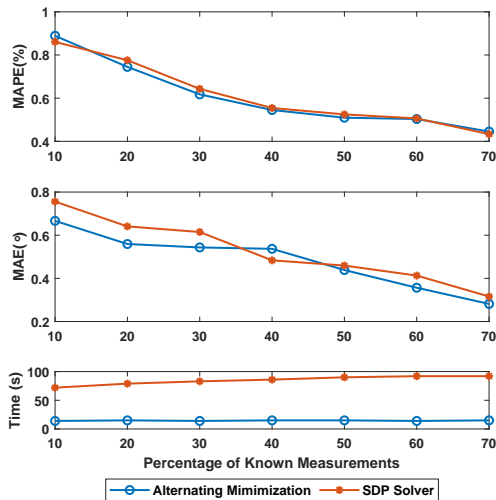


Fig. 3: Comparison of performance for the alternating minimization algorithm and SDP solver.

systems, in which buses are one-, two-, or three-phase. The slack bus is three-phase, and the total number of phases at all buses for the IEEE 123-bus system is 263 and that for the real utility feeder system is 1234. The data for both systems were simulated at 1-minute resolution using power flow analysis with diversified load and solar profiles that were created for each bus using realistic solar irradiance and load consumption data. The voltage magnitudes range from 0.95 to 1 p.u., and the voltage angles are around 0 or ± 120 degrees.

As in Section II-B, the data matrix is formulated using the real and imaginary parts of voltage phasor, voltage magnitude, and active and reactive power injection at each phase of nonslack buses. We consider the real and imaginary parts of voltage phasor as variables, and the voltage magnitude, active power, and reactive power as potentially known measurements. In our experiments, we define the percentage of known measurements as the ratio of the number of known measurements and the total number of voltage magnitude, active power, and reactive power in the data matrix. If less than two thirds of the measurements are known, there are

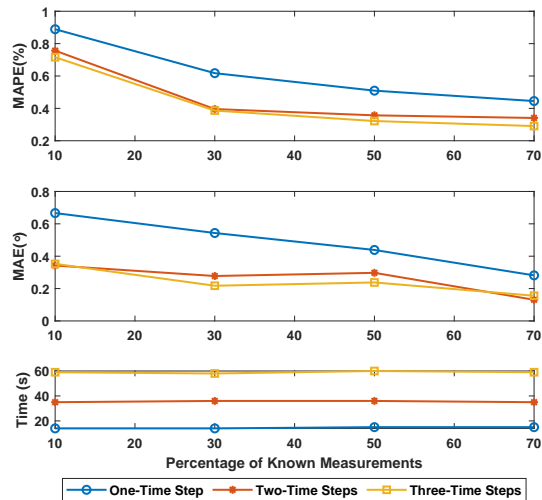


Fig. 4: Comparison of performance for alternating minimization algorithm on one-time step, two-time steps, and three-time steps.

fewer measurements than variables, under which case, the system is underdetermined and the WLS cannot generate a unique solution. For all of our experiments, the known measurements are added with noise with zero mean and 1% of the true value as the standard deviation. And the simulation results are based on the average of 50 runs for each scenario.

A. Performance on IEEE 123-Bus System

First, we implement our algorithm with one-time step data formulation at one time slot when solar generations are nonzero with 10% to 70% measurements available (where WLS is under-determined). We compare the alternating minimization algorithm with the SDP solver used by [11]. Both algorithms are run with MATLAB on a laptop with 1.9-GHz CPU and 32 GB RAM.

Fig. 2 shows one representative convergence result (with 50% available measurements) of the alternating minimization algorithm. Fig. 3 shows that the mean absolute percentage errors (MAPEs) of the voltage magnitude and the mean absolute errors (MAEs) of the voltage angle are comparable for the alternating minimization algorithm and the SDP solver. The running time for the alternating minimization algorithm is about 15 seconds for different percentages of available measurements, whereas that for the SDP solver is more than 70 seconds. These results show that the alternating minimization algorithm estimates voltage phasors more computationally efficient than the SDP solver under similar performance. The running time for the SDP solver increases as more measurements are available. This is because the SDP solver solves a constrained SDP problem, where more measurements being available means more constraints to be satisfied, and thus longer time is required to achieve the convergence.

To test the computation time and estimation accuracy as more data are used, we implement the alternating minimization algorithm on the two-time step and three-time step data matrix formulation with 10%, 30%, 50%, and 70% measurements available. Fig. 4 shows that the MAPEs of the voltage magnitudes and the MAEs of the voltage angles for the two-time step and three-time step cases are much better than the one-time step case. This is because the matrix completion problem minimizes the rank, and the larger the time steps are, the smaller the rank of the matrix is compared to its size (which is shown in Fig. 1) and the more accurate the estimates are. The running time also increases with the length of time steps, however, because of the increased size of the matrices. The MAPEs of the voltage magnitudes and MAEs of the voltage angles for the three-time step case is slightly better than the two-time step case overall, but the running time almost doubles. However, the running time for the three scenarios shows that the alternating minimization algorithm for state estimation can be implemented in real time for the IEEE 123-bus system.

B. Performance on A Real Feeder System

In order to demonstrate the scalability of our algorithm, we implement it on a real utility feeder system with 1234 phases under the assumption that 50% measurements are available for one-time, two-time, and three-time step matrix formulation, respectively. The performance in terms of MAPE, MAE, and running time are shown in Table I. The results show that the performance improves with more data being used for building the matrix and the computation time also increases. However, the running time shows that the algorithm can be applicable in real-time estimation.

TABLE I: Performance on a real utility feeder system

	MAPE	MAE	Time (s)
One-time step matrix formulation	0.4840	0.3967	50
Two-time step matrix formulation	0.4240	0.2869	105
Three-time step matrix formulation	0.3503	0.2265	156

V. CONCLUSIONS

This paper applied the constrained matrix completion formulation for state estimation in multiphase distribution systems. Different from the constrained matrix completion, the data matrix was formulated using time-series data and the alternating minimization algorithm was applied to solve the constrained matrix completion. The summable convergence of the alternating minimization algorithm was proved. In addition, the efficacy and scalability of the algorithm were demonstrated via the IEEE 123-bus system and a real utility feeder system.

REFERENCES

[1] A. Abur and A. Gomez-Exposito, *Power System State Estimation: Theory and Implementation*. Abingdon: Dekker, 2004.
[2] M. Barukčić, M. Vukobratović, D. Masle, D. Buljić, and Ž. Herderić, "The evolutionary optimization approach for voltage profile estimation in a radial distribution network with a decreased number of measurements," in *International Conference on Electrical Machines, Drives and Power Systems (ELMA)*, Jun 2017, pp. 26–31.

[3] V. Rigoni and A. Keane, "Remote voltage estimation in lv feeders with local monitoring at transformer level," in *2017 IEEE Power Energy Society General Meeting*, Jul. 2017, pp. 1–5.
[4] E. Manitsas, R. Singh, B. C. Pal, and G. Strbac, "Distribution system state estimation using an artificial neural network approach for pseudo measurement modeling," *IEEE Transactions on Power Systems*, vol. 27, no. 4, pp. 1888–1896, Nov. 2012.
[5] M. Pertl, K. Heussen, O. Gehrke, and M. Rezkalla, "Voltage estimation in active distribution grids using neural networks," in *2016 IEEE Power and Energy Society General Meeting*, Jul. 2016, pp. 1–5.
[6] H. Jiang and Y. Zhang, "Short-term distribution system state forecast based on optimal synchrophasor sensor placement and extreme learning machine," in *2016 IEEE Power and Energy Society General Meeting*, Jul. 2016, pp. 1–5.
[7] A. S. Zamzam, X. Fu, and N. D. Sidiropoulos, "Data-driven learning-based optimization for distribution system state estimation," *IEEE Transactions on Power Systems*, vol. 34, no. 6, pp. 4796–4805, 2019.
[8] K. A. Clements, "The impact of pseudo-measurements on state estimator accuracy," in *2011 IEEE Power and Energy Society General Meeting*, Jul. 2011, pp. 1–4.
[9] J. Wu, Y. He, and N. Jenkins, "A robust state estimator for medium voltage distribution networks," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1008–1016, May 2013.
[10] X. Zhou, Z. Liu, Y. Guo, C. Zhao, L. Chen, and J. Huang, "Gradient-based multi-area distribution system state estimation," *arXiv preprint arXiv:1909.11266*, 2019.
[11] P. Donti, Y. Liu, A. Schmitt, A. Bernstein, R. Yang, and Y. Zhang, "Matrix completion for low-observability voltage estimation," *IEEE Transactions on Smart Grid*, vol. 11, no. 3, pp. 2520–2530, May 2020.
[12] E. Candès and B. Recht, "Exact matrix completion via convex optimization," *Communications of the ACM*, vol. 55, no. 6, pp. 111–119, Jan. 2012.
[13] P. Gao, M. Wang, S. G. Ghiocel, J. H. Chow, B. Fardanesh, and G. Stofopoulos, "Missing data recovery by exploiting low-dimensionality in power system synchrophasor measurements," *IEEE Transactions on Power Systems*, vol. 31, no. 2, pp. 1006–1013, Mar. 2016.
[14] M. Liao, D. Shi, Z. Yu, Z. Yi, Z. Wang, and Y. Xiang, "An alternating direction method of multipliers based approach for pmu data recovery," *IEEE Transactions on Smart Grid*, vol. 10, no. 4, pp. 4554–4565, Jul. 2019.
[15] C. Genes, I. Esnaola, S. M. Perlaza, L. F. Ochoa, and D. Coca, "Robust recovery of missing data in electricity distribution systems," *IEEE Transactions on Smart Grid*, vol. 10, no. 4, pp. 4057–4067, Jul. 2019.
[16] Y. Koren, R. Bell, and C. Volinsky, "Matrix factorization techniques for recommender systems," *Computer*, vol. 42, no. 8, pp. 30–37, Aug. 2009.
[17] Z. Luo and P. Tseng, "Error bounds and convergence analysis of feasible descent methods: a general approach," *Annals of Operations Research*, vol. 46, no. 1, pp. 157–178, Mar. 1993.
[18] Y. Xu and W. Yin, "A block coordinate descent method for multi-convex optimization with applications to nonnegative tensor factorization and completion," *SIAM J. Imaging Sci.*, vol. 6, no. 3, pp. 1758–1789, Sep. 2013.
[19] A. Bernstein, C. Wang, E. Dallanese, J.-Y. L. Boudec, and C. Zhao, "Load-flow in multiphase distribution networks: Existence, uniqueness, non-singularity, and linear models," *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 5832–5843, Nov. 2018.
[20] K. Christakou, J. LeBoudec, M. Paolone, and D. Tomozei, "Efficient computation of sensitivity coefficients of node voltages and line currents in unbalanced radial electrical distribution networks," *IEEE Transactions on Smart Grid*, vol. 4, no. 2, pp. 741–750, Jun 2013.
[21] S. S. Guggilam, E. Dall'Anese, Y. C. Chen, S. V. Dhople, and G. B. Giannakis, "Scalable optimization methods for distribution networks with high pv integration," *IEEE Transactions on Smart Grid*, vol. 7, no. 4, pp. 2061–2070, Jul. 2016.
[22] N. Srebro, *Learning with Matrix Factorizations*. Ph.D. thesis, Massachusetts Institute of Technology, 2004.
[23] P. Jain, P. Netrapalli, and S. Sanghavi, "Low-rank matrix completion using alternating minimization," in *Annual ACM Symposium on Symposium on Theory of Computing - STOC 13*, 2013.
[24] J. Lambers, "Coercive functions and global minimizers," <https://www.math.usm.edu/lambers/mat419/lecture4.pdf>, 2011.
[25] IEEE, "Resources: 123-bus feeder," <http://sites.ieee.org/pes-testfeeders/files/2017/08/feeder123.zip>.