An Alternating Direction Method of Multipliers Based Approach for PMU Data Recovery

Mang Liao, Student Member, IEEE, Di Shi, Senior Member, IEEE, Zhe Yu, Member, IEEE, Zhehan Yi, Member, IEEE, Zhiwei Wang, Member, IEEE, and Yingmeng Xiang, Student Member, IEEE

Abstract—This paper presents a novel algorithm for recovering missing phasor measurement unit (PMU) data. Due to the low-rank property of PMU data, missing measurement recovery can be formulated as a low-rank matrix-completion problem. Based on maximum-margin matrix factorization, we propose an efficient algorithm, alternating direction method of multipliers (ADMM), for solving the matrix completion problem. Compared to existing approaches, the proposed ADMM based algorithm does not need to estimate the rank of the target data matrix and provides better performance in computation complexity. Since PMU data are transferred through insecure and delayed communications, we consider the case of measurements missing from all PMU channels in several sampling instants and provide the strategies for reshaping the matrix composed by the received PMU data for the recovery. Numerical results using real PMU measurements from State Grid Jiangsu Electric Power Company illustrate the effectiveness and efficiency of the proposed approaches.

I. INTRODUCTION

The wide-area measurement system (WAMS) using phasor measurement units (PMUs) has been regarded as one of the key enabling technologies in monitoring, control, and protections of the next-generation power grids [1]. With continuous increase in PMU deployment and the resulting explosion in data volume, the design and deployment of an efficient wide-area communication and computing infrastructure, especially from the point of view of resilience against a large number of missing data, is evolving as one of the greatest challenges to the power system and IT communities. With thousands of networked PMUs being scheduled to be installed in the United States by 2020, exchange of synchrophasor data between balancing authorities for any type of wide-area control will involve an enormous number of data flow in real-time per event, thereby opening up a wide spectrum of probabilities of data losses and data quality degradation in an unpredictable way. Data missing makes the system less observable, degrades the performance of the state estimates and system parameter identification and verification [2], and weakens the security and stability of the system. Therefore, recovering missing PMU measurements has become a significant and inevitable problem in power systems.

Classical approaches of missing data recovery are conducted on an individual channel basis, without considering the cross correlation among different channels [3–5]. It has been shown that such approaches, linear interpolation for instance, can generate considerable errors or even become incapable when missing data occurs continuously over a long period [6–8]. In practice, it is very common that PMU data are missing over a long period and data recovery becomes critical to many applications. For example, during a Florida event in 2008, PMU measurements in the Eastern Interconnection (EI) experience missing data problem over multiple continuous sampling instants [9]. Increasing deployment of PMUs allows simultaneous analysis of PMU measurements from multiple channels. PMU data can be structured as a matrix with each column and row representing measurements from one channel and sampling instant, respectively. Since large amounts of PMU data exhibit highly correlated property [10–12], the matrix is approximately low-rank. In power systems, such blocks of PMU data have been used to study the propagation of oscillations after a disturbance [13, 14], and to detect the system stability using a singular value approach [15].

The problem of recovering the missing PMU data can be formulated as a low-rank matrix-completion problem [6, 11].
Studies on matrix completion algorithms are extensive, including atomic decomposition for minimum-rank approximation (ADMiRa) [16], singular value projection (SVP) [17], information cascading matrix completion (ICMC) [18], among which nuclear-norm-regularized matrix approximation [19–22] and maximum-margin matrix factorization (MMMF) [23] are widely adapted. Using nuclear-norm-regularized matrix approximation, a singular value threshold (SVT) has to be designed which influences the recovery accuracy. In addition, the singular value decomposition (SVD) calculation increases the computational time and complexity. Based on MMMF, alternating least squares (ALS) schemes are proposed for solving the matrix-completion problem [11, 24–27]. Further, a software package softImpute is developed for reducing the computation time [28].

Most of the existing approaches rely on an estimation of the rank $r$ of the data matrix, which is typically unavailable and time variant in practice. Inaccurate estimation of $r$ introduces modeling errors in the matrix completion problem. In the update of iterates, ALS has to calculate the inverse of an $r$-by-$r$ matrix, and the computational complexity is proportional to $r^3$. On the other hand, $r$ cannot be too small for guaranteeing recovery accuracy. Therefore, design of an adaptive and scalable online algorithm of PMU data recovery is an open challenge.

The communication between PMUs and phasor data concentrators (PDCs) often suffers natural and artificial disturbances, resulting in delayed or distorted data. In this case all elements in one or more rows and columns of the PMU data matrix may be missing, and they cannot be recovered using existing methods. Recovering the lost data caused by malfunction of PMUs is still an unresolved problem.

Motivated by these insights, we develop an algorithm that can recover the missing PMU measurements with lower computational complexity and less computation time. The fundamental set-up for this optimization was based on MMMF and alternating direction method of multipliers (ADMM) [31–33]. Firstly, the PMU measurements with possible data missing are structured as a matrix $M \in \mathbb{R}^{n_1 \times n_2}$ whose columns and rows represent the measurements from one channel and the same sampling instant, respectively. Secondly, we formulate the data recovery as an optimization problem in which we minimize the rank of the recovery matrix $\hat{X}$ while keeping elements in $\hat{X}$ the same as the corresponding ones in $M$ if they are present. Then an ADMM algorithm is proposed to solve the optimization problem in an iterative way. In the update equations there is no matrix inverse computation, which immensely reduces the computational complexity. Due to the less computation time, ADMM algorithm for the recovery can be used for off-line applications such as post event analysis and system identification, as well as real-time data-driven analytic such as state estimation and disturbance identification.

In addition, it is not necessary to estimate the rank of the ideal data matrix $X$ without missing data, which significantly cuts down the influence of the uncertain factor into the performance.

Furthermore, we consider the case of missing data from all PMU channels. In this case, all elements in one or multiple rows and columns of the observed matrix $M$ are missing. Hankel matrix is used for recovering the lost data from all channels in [29]. However, rank of the filled matrix has to be estimated for the recovery. Preliminary results on this problem for recovering the missing PMU data from all channels at one sampling instant have been reported in our recent paper [30]. This paper expands the analysis to a more general case where PMU data are missing from all channels in multiple sampling instants. Two efficient algorithms are presented to reshape the observed matrix, and the lost data from all channels during several sampling instants can be recovered using the ADMM approach. ADMM utilizes the spatial and temporal correlations of the PMU data for the recovery. If the missing data are correlated to the observed data in temporal or spatial dimension, it can estimate the measurements successfully. Thus, combining ADMM and the approach of reshaping the observed matrix, the recovery can be applied in the cases, such as lost data on one channel during multiple successive sampling instants, lost data in one sampling instant on multiple channels, and lost data on all the channels during one or multiple sampling instants in both steady-state and disturbance scenarios. Numerical experiments on actual PMU data collected from State Grid Jiangsu Electric Power Company are conducted to verify the effectiveness of the proposed methods.

The main contributions of this paper are as follows: (1) We provide a more efficient algorithm, ADMM, to recover the missing PMU measurements. It makes lower computational complexity, and avoids to estimate the rank of the filled complete matrix, which reduces the influence of the uncertain factor into the performance. Due to the smaller computation time of the recovery, the ADMM algorithm can be used for both off-line and online applications. (2) We proposed a set of algorithms to recover the missing PMU data from all the channels during several sampling instants. These data cannot be recovered using existing approaches. The methods of restructuring the PMU data are presented, and combining to the ADMM approach the missing measurements can be recovered if one or more PMUs are failed in multiple sampling instants. (3) Compared to the interpolation methods, ADMM algorithms can recover the lost data on multiple channels during successive several sampling instants based on the spatial and temporal correlations of PMU data. ADMM expands the applicable scenarios. (4) We illustrate our results with actual PMU measurements from Jiangsu Electric Power Company in China. The computation time of recovery using ADMM is around 4s with 400 iterations for a 2800-by-72 observed matrix, which satisfies the requirement of real-time state estimation. The computation time is calculated using Matlab with a 1.6GHz processor and an 8GB memory.

The remainder of this paper is organized as follows. Section II represents the low-rank property of PMU measurements and formulates the recovery problem as a low-rank matrix completion problem. Section III provides an ADMM based approach with less computational complexity, and section IV presents the reshaping methods for recovering the lost data from all channels in one or multiple sampling instants. Section V envisions the applications using ADMM for the
recovery. Section VI verifies algorithms using actual PMU measurements from Jiangsu Electric Power Company and section VII concludes the paper.

II. PROBLEM FORMULATION

Interpolation, such as linear interpolation and spline interpolation, is a simple and traditional method to recover the missing PMU data. It utilizes the temporal correlation to recover the lost data in one channel. If missing data are sparse or the data at the previous and later sampling instants of the missing data are observed, the recovery accuracy and computation time using interpolation method satisfy the requirements of the real-time applications. However, if the measurements in one channel are missing during a long period, using interpolation method for the recovery is not an advisable choice. In this section, we process a spatial-temporal block of PMU data, present the low-rank property of PMU data, and formulate the data recovery as a matrix completion problem.

A. Low-rank property of PMU measurements

Denote \( X = [x_1, x_2, \ldots, x_{n_2}] \in \mathbb{R}^{n_1 \times n_2} \) as the ideal PMU measurement matrix without data missing. Each \( n_1 \)-by-1 vector \( x_i \), \( 1 \leq i \leq n_2 \), denotes the measurements on the \( i^{th} \) channel during \( n_1 \) sampling instants. The structure of matrix \( X \) only depends on the time-varying PMU data. Consider an example to illustrate the structure of the matrix containing PMU data. We construct a 2800-by-72 matrix \( X_1 \) using oscillation event measurements with DC component removed from Jiangsu Electric Power Company. Three phase voltage magnitude measurements for 112s from 24 buses are placed side by side to form the ideal data matrix \( X_1 \) without missing data. The sample rate is 25Hz. A 1025-by-63 matrix \( X_2 \) is constructed similarly using voltage magnitude measurements for 41s from 21 buses in an ambient scenario. From Fig. 1, it can be seen that the PMU data are strongly correlated in both temporal and spatial dimensions, which implies that \( X_1 \) and \( X_2 \) are strongly row and column correlated, respectively.

Ignoring the noise, since the PMU measurements of voltage or current phasors or magnitudes on different channels are strongly correlated [10–12], the matrix containing PMU data is a low rank matrix, i.e., most of its singular values are equal to zero. However, the actual PMU data are noisy, which weakens the correlation. Only a few singular values of the ideal data matrix \( X \) is significant. Most of them are close to zero. Fig. 2 shows the singular values of \( X_1 \) and \( X_2 \), respectively. Only few singular values are much larger than the others in both \( X_1 \) and \( X_2 \) matrices, which illustrates that \( X_1 \) and \( X_2 \) are approximately low-rank.

An approximating rank approach, referred to as Frobenius norm proportion [34], is stated as follows.

\[
\sqrt{\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_r^2} \leq \beta \sqrt{\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_r^2 + \cdots + \sigma_l^2} 
\]

where \( \sigma_1 > \sigma_2 > \cdots > \sigma_l > 0 \) are the singular values of the matrix and \( \beta, 0 < \beta \leq 1 \), is the proportion factor. \( r \) in (1) denotes the approximate rank of the matrix. If \( \beta = 1 \), the approximate rank of \( X \) is equal to \( l \). Since the PMU measurements of voltage or current phasors or magnitudes from different lines or buses are strongly correlated [10–12], the approximate rank of \( X \) is much smaller than \( \min \{ n_1, n_2 \} \). If the proportion factor \( \beta = 0.993 \), the approximate ranks of \( X_1 \) and \( X_2 \) as in the example of Fig. 1 are 4 and 6, respectively.

Now we consider the influences of PMU data sampling rate and the length of data window into the approximate rank of the matrix containing the PMU data. For fairness, the value of \( \beta = 0.993 \) is fixed with different data sampling rates and lengths of data windows. Fig. 3 shows the approximate ranks against different lengths of data windows in both dynamic and steady-state scenarios. In both scenarios, the approximate ranks are no more than 6. That illustrates that with different and long enough lengths of data windows (\( t \geq 1s \)), matrices \( X \) are approximately low-rank. Fig. 4 shows the approximate ranks against different data rates in both dynamic and steady-state scenarios. With rates 150/min, 214/min, 375/min, 750/min, and 1500/min, the approximate ranks are 4 and 6 in dynamic and steady-state scenarios, respectively. It illustrates that with different rates, matrices with PMU data are approximately low-rank, and the influence of the data rate into the approximate rate can be ignored. In addition, in [10–12] the matrix containing PMU data with rate 1800/min is approximately low-rank.
where the nuclear norm \( ||\hat{X}||_* \) is the sum of the singular values of \( \hat{X} \).

In [23], the authors propose MMMF to solve (4) as

\[
\min_{A,B} \frac{1}{2}(||A||_F^2 + ||B||_F^2) \\
\text{subject to } (A^T B - M) \odot I_s = 0,
\]

(5)

where a \( z \)-by-\( n_1 \) matrix \( A \) and a \( z \)-by-\( n_2 \) matrix \( B \), \( z \leq \min\{n_1, n_2\} \), are introduced in which \( \hat{X} = A^T B \). Note that there is no decomposition rule for designing matrices \( A \) and \( B \). For simple computation, in the paper we let \( \hat{X} = A^T \), \( B = I \), and \( z = n_2 \), where \( I \) is an identity matrix. The matrix completion problem in equation (5) is a bi-convex problem. Fixing matrix \( A \), the optimization is convex of matrix \( B \), and vice versa. Iterative algorithms including ALS have been proposed to solve (5). In this work, we propose an ADMM based approach, which does not require estimating the approximate rank of the measurement matrix \( X \).

III. AN ADMM BASED APPROACH FOR PMU DATA RECOVERY

Since the objective function in (5) is a convex function, the ADMM method can be applied to solve the optimization problem (5) in an iterative way using the Lagrangian multiplier approach [31]. The augmented Lagrangian for (5) can be formulated as

\[
\mathcal{L} = \frac{1}{2}(||A||_F^2 + ||B||_F^2) + \text{trace}(w^T((A^T B - M) \odot I_s)) + \frac{\rho}{2}||((A^T B - M) \odot I_s)||_F^2,
\]

(6)

where \( A \) and \( B \) are the matrices of the primal variables, \( w \in \mathbb{R}^{n_1 \times n_2} \) is the matrix of the dual variables, and penalty factor \( \rho > 0 \) denotes the step size of the dual variable update. The ADMM update equations suggest that large values of \( \rho \) place a large penalty on violations of primal feasibility and so tend to produce small primal residuals. On the other hand, small values of \( \rho \) tend to reduce the dual residual, but may result in a larger primal residual. A scheme of designing a time-varying penalty weight based on the comparison between the norm of primal variable and dual variable is provided in [38]. For reducing the computation time, we consider \( \rho \) as a constant number. Under the condition of guaranteeing the convergence, we choose a large value of \( \rho \) for fast convergence rate. Our numerical experiments in Section VI suggest that \( \rho = 10^{-6} \).

Let \( \langle A, B \rangle \) denote the standard inner product of matrices \( A \) and \( B \) where \( \langle A, B \rangle = \sum_{i,j} [A]_{ij} [B]_{ij} \). It can be shown that \( \langle M_1, M_2 \rangle := \text{trace}(M_1^T M_2) \) [34]. Thus we have

\[
\text{trace}(w^T((A^T B - M) \odot I_s)) = \langle w, (A^T B - M) \odot I_s \rangle \geq \langle w \odot I_s, A^T B - M \rangle = \text{trace}((w \odot I_s)^T (A^T B - M));
\]

\[
||(A^T B - M) \odot I_s||_F^2 = \langle (A^T B - M) \odot I_s, (A^T B - M) \odot I_s \rangle \geq \langle (A^T B - M) \odot I_s, (A^T B - M) \odot I_s \rangle = \text{trace}(((A^T B - M) \odot I_s)^T (A^T B - M)).
\]
The augmented Lagrangian (6) can be rewritten as
\[
\mathcal{L} = \frac{1}{2}(||A||_F^2 + ||B||_F^2) + \text{trace}[(w \odot I_s) \mathbf{T}^T(A^T B - M)] + \frac{\rho}{2}\text{trace}[(A^TB - M) \odot I_s]^T(A^TB - M).
\]
(7)

The gradients of \( \mathcal{L} \) in (7) with respect to \( A \) and \( B \) are stated as follows.
\[
\frac{\partial \mathcal{L}}{\partial A} = A + B(w \odot I_s)^T + \rho B[(A^TB - M) \odot I_s]^T,
\]
\[
\frac{\partial \mathcal{L}}{\partial B} = B + A(w \odot I_s) + \rho A(A^TB - M) \odot I_s.
\]
(8)

Given the derivation, the ADMM algorithm for solving the optimal problem (5) is illustrated in Algorithm 1. The parameter \( \epsilon \) in the stopping criterion of Algorithm 1 influences the iteration number and recovery accuracy. In general, with smaller value of \( \epsilon \), the iteration number is increased and the recovery accuracy is enhanced. Note that since using ADMM algorithm the initial values do not influence the convergence [40], the initial values of \( z \)-by-\( n_1 \) \( A \) and \( n_2 \)-by-\( z \) \( B \), \( z \leq \min\{n_1, n_2\} \), can be any matrix which satisfies \( \hat{X}^0 = (A^0)^T B^0 \). For simple expression, we let \( (A^0)^T = \hat{X}^0 \), \( B^0 = I \), and \( z = n_2 \).

Algorithm 1 ADMM algorithm for PMU data recovery

Initialize \( \hat{X}^0 = M \), \( (A^0)^T = \hat{X}^0 \), \( B^0 = I \), and \( w^0 \), and determine the value of \( \rho \), \( \epsilon \), and \( k_{\text{max}} \).

Do:
\[
A^{k+1} = -B^k(w^k \odot I_s)^T - \rho B^k[(A^{k+1}B^{k+1} - M) \odot I_s]^T,
\]
\[
B^{k+1} = A^k(w^k \odot I_s) + \rho(A^{k+1}B^{k+1} - M) \odot I_s,
\]
\[
w^{k+1} = w^k + \rho[(A^{k+1})^T B^{k+1} - M] \odot I_s,
\]
\[ k = k + 1. \]

until: The stopping criterion \( ||(A^{k+1})^T B^{k+1} - (A^k)^T B^k|| < \epsilon \) is reached or \( k > k_{\text{max}} \), where \( k_{\text{max}} \) is the predetermined maximum iteration number.

The updates in Algorithm 1 requires no matrix inverse, and the computational complexity is \( \mathcal{O}(n_1 n_2 z) \). If the rank of the original data matrix \( X \) is known as \( r \), we can set \( z = r \), and the computational complexity using ADMM is \( \mathcal{O}(n_1 n_2 r) \). For avoiding to estimate the rank, we let \( z = \min\{n_1, n_2\} \), which reduces the influence of the uncertain factor into the performance. The computational complexity is \( \mathcal{O}(n_1 n_2 \min\{n_1, n_2\}) \), which is a quadratic function of \( \min\{n_1, n_2\} \).

The proposed approach utilizes the spatial-temporal correlation of the PMU data. Its applications are broad in transmission and distribution systems, with different voltage levels, across different PMU classes, in dynamic and stationary scenarios, and with different lengths of data windows \( (t \geq 1 \text{s}) \) and different data rates, whereas the measurement data have strong correlations. The performance of the proposal compared to other recovery methods, such as ALS approach, linear interpolation, and spline interpolation, in both steady-state and disturbance scenarios is examined in Section VI.

IV. MATRIX RESHAPE

Power systems often suffer natural and artificial disturbances during operations. It is possible that the data from all channels are missing simultaneously during multiple sampling instants under communication failure. In this case, no existing algorithms can recover the missing data. For solving this problem, the observed matrix \( M \) has to be reshaped to avoid its columns or rows with all zero elements. Our goal is to recover the missing data in one row of the reshaped observed matrix \( M \) as small as possible. Meanwhile the corresponding reshaped recovery matrix \( \hat{X} \) is still low-rank. We ignore the case of missing elements in one column of matrix \( M \). It can be avoided by taking more PMU measurements in a longer time. For example, one channel is failed in the first second and recovered in the next second. If \( M \) contains the PMU measurements in the first second, one column of \( M \) will own all zero elements; while if \( M \) contains the PMU measurements in the first two seconds, there is no column with all zero elements.

A. Cut-Column Reshaping Method

Considering the case of only one row of matrix \( M \) with all zero elements, we provide an alternative method for reshaping the observed matrix, called cut-column reshaping method (CCRM). Using CCRM, each column with \( n_1 \) length is separated into \( n^* \) shorter columns with a length of \( \frac{n_1}{n^*} \). Thus, the \( n_1 \)-by-\( n_2 \) matrix is reshaped to an \( \frac{n_1}{n^*} \)-by-\( n^*_2 \) matrix, and the original column correlation is held. The length of the new column should be larger than the row length of the original matrix, i.e., \( \frac{n_1}{n^*} > n_2 \). \( n^* \) also need to satisfy that \( \left\lfloor \frac{n_1}{n^* + 1} \right\rfloor < n_2 \), where \( \lfloor x \rfloor \) denotes the smallest integer number which is larger than \( x \). Thus using CCRM the numbers of rows and columns of reshaped matrix are both larger than \( n_2 \).

In addition, with holding the column correlation, CCRM minimizes the proportion of zero elements in one row of the
reshaped matrix. Consider a simple example to illustrate the reshaping method. A 6-by-2 matrix $M$ can be expressed as:

$$
M = \begin{bmatrix}
  m_1 & m_2 \\
  m_3 & m_4 \\
  m_5 & m_6
\end{bmatrix} = \begin{bmatrix}
m_{11} & m_{21} & m_{31} & m_{41} & m_{51} & m_{61} \\
m_{12} & m_{22} & m_{32} & m_{42} & m_{52} & m_{62}
\end{bmatrix}^T,
$$

(9)

whose fifth row owns all zero elements. Using CCRM, set $n^* = 3$, and matrix $M$ is reshaped into a 2-by-6 matrix:

$$
M_{re} = \begin{bmatrix}
m_{re1} & m_{re2} & m_{re3} & m_{re4} & m_{re5} & m_{re6} \\
m_{re1} & m_{re1} & m_{re1} & m_{re1} & m_{re1} & m_{re1}
\end{bmatrix}.
$$

Now the elements in any row or column are not all equal to zero. If $m_1$ and $m_2$ are strongly correlated, then using OCCRM is illustrated in Algorithm 3.

**Algorithm 3 Order- and Cut-Column Reshaping Method**

1) Check whether any row or column of the $n_1$-by-$n_2$ observed matrix $M$ owns all zero elements.
2) If yes, build up a set $S$ which contains all the divisors of $n_1$ in a descending order, whose corresponding quotients are equal or larger than $n_2$.
3) Let $n^*$ be the first entry in $S$.
4) Separate each column of $M$ into $n^*$ shorter columns with $\frac{n_2}{n^*}$ length. The original $n_1$-by-$n_2$ matrix is reshaped into a $\frac{n_1}{n^*}$-by-$n_2n^*$ matrix.
5) Check whether any row or column of the reshaped $\frac{n_1}{n^*}$-by-$n_2n^*$ matrix owns all zero elements.
6) If yes, re-arrange the row order of matrix $M$ and repeat Step 4.
7) If the reshaped matrix still has the column or row with all zero elements, let $n^*$ be the next entry in $S$ and repeat Step 4.

**Corollary 1.** Using OCCRM at least one row or column of reshaped $\frac{n_1}{n^*}$-by-$n_2n^*$ matrix is still with all zero elements, if the number of the missing rows of the original matrix is larger than

$$
\#row\text{missing} = \begin{cases}
  n_1 - \frac{n_1}{n^*}, & n_1 - \frac{n_1}{n^*} \geq \left(\frac{n_1}{n^*}\right)^2 \\
  n_1 - \frac{n_1}{n^*}, & n_1 - \frac{n_1}{n^*} < \left(\frac{n_1}{n^*}\right)^2
\end{cases},
$$

(12)

where $[x]$ denotes the largest integer number which is smaller than $x$.

**V. ENVISION THE FEASIBLE APPLICATIONS**

Though low-rank matrix completion methods are block-processing methods and have to wait $n_1$ sampling instants for the first block, the equilibrium values of matrices $A$ and $B$ in the previous block can be treated as the initial values in the next block, which can reduce the convergence iterations. As mentioned before, with different lengths of data windows ($t \geq 1s$) and data rates, matrix $X$ containing PMU data is still approximately low-rank. So the missing data can be recovered using ADMM approach with different sampling instants ($t \geq 1s$) and data rates. Due to avoiding the matrix inverse computation in the updates, the computation time in
one iteration is transitory using ADMM. Because of less computational complexity and computation time, the recovery using ADMM can be used for off-line applications such as post event analysis and system identification, as well as real-time data-driven analytic such as state estimation and disturbance identification. In Section VI, the missing data can be recovered using ADMM around 4s for a 2800-by-72 observed matrix with actual voltage magnitude measurements in 400 iterations. The numerical experiments on actual PMU data are collected from State Grid Jiangsu Electric Power Company. We use Matlab for calculating the computation time with a 1.6GHz processor and a 8GB memory. The computation time of recovering missing PMU data using ADMM satisfies the requirement of real-time state estimation.

For a large power grid, there can be multiple PMUs in one zone, and the column number of the observed matrix \( M \) can be huge. However, we can control the sampling instants of the observed PMU data to be small. Since the computational complexity of ADMM for the recovery is \( O(n_1n_2z) \), \( z = \min\{n_1, n_2\} \), the computational time will not increase much. In addition, we can limit the number of channels to be included in the \( M \) matrix, without affecting the strong spatial and temporal correlation among the data. With an appropriate size of the observed matrix, the computation time of the recovery using ADMM can satisfy the requirement of online applications such as state estimation and disturbance identification. Moreover, the recovery can be carried out by each PMU using individual data in a parallel manner.

Based on the spatial and temporal correlations of PMU data, estimating the missing measurements using ADMM approach can be used in the following cases: lost data in one channel during a long period, lost data in one sampling instant on several channels, and lost data in multiple sampling instants on multiple channels. In addition, combining OCCRM approach, it is possible to use ADMM algorithm for estimating the missing data on all the channels during one or multiple sampling instants. Due to the transitory computation time, acceptable estimate accuracy, and various applicable scenarios, before processing and analyzing the PMU data, the PDC can estimate PMU measurements using combination of ADMM and OCCRM algorithms first. In reality, a computer connected to the output of the PDC provides the users with software such as real time dynamics monitoring system (RTDMS) that calculates and displays locally measured frequencies, primary voltages, currents, MWs and MVARs for system operators. We can add one block before the calculation in RTDMS for recovering the missing data using the combination of ADMM and OCCRM. Then after operating all the softwares, the PDC transmits the collected measurements to synchrophasor platform and energy management system (EMS) e-terraplatform. According to the integrate of measurement-based and model-based model [39], Fig. 5 shows the feasible applications of using ADMM for the recovery.

**VI. NUMERICAL RESULTS**

To verify the proposed algorithms, we present the numerical experiments on the PMU datasets from Jiangsu Electric Power Company, which is one of the largest provincial power company in China. It has generation capacity of 100GW and peak load of 92GW. Over 160 PMUs, with thousands of measurement channels, have been installed in the system. These PMUs cover all 500kV substations, a majority of the 220kV substations, major power plants, and all renewable power plants.

In this section, we build up two PMU data matrices \( X_1 \) and \( X_2 \) whose columns and rows correspond to a real measurement sequence of voltage magnitudes (V) without DC component on 24 buses in disturbance scenario and 21 buses in stationary scenario, and the sampling instant, respectively. The measurements are observed during 112s and 41s, respectively. There are 25 samples in one second. The 2800-by-72 matrix \( X_1 \) and the 1025-by-63 matrix \( X_2 \) are full matrices with no missing measurements. If the proportion factor \( \beta = 0.993 \), the approximate ranks of \( X_1 \) and \( X_2 \) are 4 and 6, respectively.

To test the recovery accuracy of the presented ADMM algorithm, some of the observed data in \( X_1 \) and \( X_2 \) are set to be lost. Since the PMU data are missing arbitrary and unpredictable, in this paper we consider three cases of missing data: (1) Missing data randomly. The delivery of PMU measurements from multiple remote locations of power grids to monitoring centers can result in the random unavailability of PMU measurements; (2) Missing data in one channel during one second. If one channel is delayed or suffers denial-of-service (DoS) attack, then the data through that channel during the time will be lost; (3) Missing data in all channels during one second. If the communication of PMU data are failed in one second, then the 25 continuous rows of observed matrix that contains PMU measurements own all zero elements. The PMU measurements obtained from the transmission part of Jiangsu Electric Power Company are in P class. PMUs are required at sub-stations and several places in the network including tap-changing transformers, complex loads, and PV generation buses. Since in the same local area the PMU data at any place of network are strongly correlated in both P class and M class, using ADMM the missing data in M class at other parts of the power system can also be recovered.

In disturbance scenario a low frequency oscillation event
is observed by PMUs. The disturbance time is much longer than the time duration of data missing (one second). In this paper, the computational time is obtained by operating Matlab programming with a 1.6GHz processor and a 8GB memory.

A. Benchmark: Alternating Least Squares Algorithms

A matrix completion method, referred to as alternating least squares (ALS) provided in [24, 25], is employed as a benchmark in the simulation. It assumes the rank of $\hat{X}$ is known as $r$, $r \ll \min\{n_1, n_2\}$, and set $z = r$. Thus, $A = \{a_1, a_2, \ldots, a_{n_1}\} \in \mathbb{R}^{r \times n_1}$ and $B = \{b_1, b_2, \ldots, b_{n_2}\} \in \mathbb{R}^{r \times n_2}$. The objective function in (5) can be rewritten as

$$
\min_{A,B} \sum_{i,j:[I_2]_{ij}=1} ([\hat{X}]_{ij} - a_i^T b_j)^2 + \lambda (||A||^2_F + ||B||^2_F), (13)
$$

where $\lambda > 0$ is the dual parameter. The ALS algorithm solves the optimization problem in (13) in an iterative way. For $1 \leq i \leq n_1$, $1 \leq j \leq n_2$, the updates at iteration $k + 1$ are

$$
a_{i}^{k+1} = (\sum_{j:[I_2]_{ij}=1} b_j^{(k)}(b_j^{(k)})^T + \lambda I_r)^{-1}(\sum_{j:[I_2]_{ij}=1} [\hat{X}]_{ij} b_j^{(k)}),
$$

$$
b_{j}^{k+1} = (\sum_{i:[I_2]_{ij}=1} a_i^{(k)}(a_i^{(k)})^T + \lambda I_r)^{-1}(\sum_{i:[I_2]_{ij}=1} [\hat{X}]_{ij} a_i^{(k)}).
$$

The computational complexity of ALS algorithm is $O(r^3(n_1 + n_2))$. It can be reduced by smaller $r$. However, $r$ is an estimated parameter which is not available in practice and cannot be too small to guarantee the recovery accuracy, especially if the number of erased data is large. The convergent rate depends on the dual parameter $\lambda$. Large $\lambda$ makes convergent rate fast, but reduces the recovery accuracy. Thus, it is a trade-off problem for selecting $r$ and $\lambda$ in ALS algorithm.

For recovering the missing PMU data, we choose the penalty factor $\rho = 10^{-6}$ using ADMM, and $\lambda = 5$ and $r = 20$ using ALS. Also we choose $\epsilon = 100$ and $k_{max} = 500$. Since the recovery accuracy is not improved a lot with a smaller value of $\epsilon$, we choose a larger value of $\epsilon$ for reducing the computational time.

B. Case 1: Random Miss

In this case, we assume an independent and identical distribution (i.i.d.) of the missing rate. For each data point in $X$, with a probability of $1 - p$, the measurement is missing in $M$.

Fig. 6 shows the trajectories of the first elements of the recovered matrices $X_1$ and $X_2$ using ADMM and ALS, respectively. Using ALS the recovery matrices converges within 22 and 28 iterations with voltage magnitude measurements in disturbance and stationary scenarios, respectively. Using ADMM the recovery matrices converge within 199 and 230 iterations. Furthermore, during first several iterations, the estimated data are oscillating using ADMM; while using ALS the estimated data are convergent, which implies that the convergence of ALS outperforms ADMM.

Fig. 7 shows the computational time using ADMM and ALS with voltage magnitude measurements in disturbance and stationary scenarios, respectively. In both scenarios, the computational time using ADMM is less than 2s. Though using ADMM the convergent iteration number is larger than the one of ALS, the computational time is much smaller. This is because the computational time in one iteration using ALS is much longer due to matrix inverse computation.

Fig. 8 shows the statistic relative recovery errors (RREs) using ADMM and ALS with voltage magnitude measurements in disturbed and stationary scenarios against different observed data probabilities ($\sum[I_2_{ij}] / (n_1 \times n_2)$, respectively. The expression of RRE is given by

$$
\text{RRE} = \frac{||[\hat{X}_2 - X_2] \odot (1_{n_1 \times n_2} - I_{s})||_F}{||X_2 \odot (1_{n_1 \times n_2} - I_{s})||_F}, (14)
$$

where $1_{n_1 \times n_2}$ is an $n_1$-by-$n_2$ matrix composed by 1s. The statistic values of RREs are calculated using Monte Carlo runs. With smaller missing rate, RREs become smaller. In disturbance scenario, at each observed data probability, RREs using ADMM are smaller than using ALS. Compared to the voltage magnitude measurements in the stationary scenario, RRE is lower in disturbance scenario, since the approximate rank of $X_1$ is smaller than $X_2$. Tables I and II list the statistic values of convergence iterations, computational time, and RREs with different observed data probabilities in disturbance and stationary scenarios using ADMM, ALS, linear interpolation, and spline interpolation.
rerespectively. Here $p$ denotes the observation probability of data.

### Case 1: Comparison among ALS, ADMM, Linear, and Spline interpolations for the recovery in disturbance scenario

<table>
<thead>
<tr>
<th></th>
<th>ADMM</th>
<th>ALS</th>
<th>Linear</th>
<th>Spline</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p = 0.9$</td>
<td># iteration</td>
<td>time (s)</td>
<td># iteration</td>
<td>time (s)</td>
</tr>
<tr>
<td></td>
<td>0.1344</td>
<td>0.1646</td>
<td>0.0979</td>
<td>0.1355</td>
</tr>
<tr>
<td>$p = 0.8$</td>
<td># iteration</td>
<td>time (s)</td>
<td>0.65</td>
<td>0.75</td>
</tr>
<tr>
<td>$p = 0.7$</td>
<td># iteration</td>
<td>time (s)</td>
<td>0.1633</td>
<td>0.26</td>
</tr>
<tr>
<td>$p = 0.6$</td>
<td># iteration</td>
<td>time (s)</td>
<td>0.1721</td>
<td>0.2674</td>
</tr>
<tr>
<td>$p = 0.5$</td>
<td># iteration</td>
<td>time (s)</td>
<td>0.1765</td>
<td>0.3535</td>
</tr>
</tbody>
</table>

### Case 2: Comparison among ALS, ADMM, Linear, and Spline interpolations for the recovery in stationary scenario

<table>
<thead>
<tr>
<th></th>
<th>ADMM</th>
<th>ALS</th>
<th>Linear</th>
<th>Spline</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p = 0.9$</td>
<td># iteration</td>
<td>time (s)</td>
<td>1.3</td>
<td>1</td>
</tr>
<tr>
<td>$p = 0.8$</td>
<td># iteration</td>
<td>time (s)</td>
<td>1.229</td>
<td>1.187</td>
</tr>
<tr>
<td>$p = 0.5$</td>
<td># iteration</td>
<td>time (s)</td>
<td>1.39</td>
<td>1.15</td>
</tr>
<tr>
<td>$p = 0.6$</td>
<td># iteration</td>
<td>time (s)</td>
<td>1.312</td>
<td>1</td>
</tr>
<tr>
<td>$p = 0.5$</td>
<td># iteration</td>
<td>time (s)</td>
<td>1.39</td>
<td>10</td>
</tr>
</tbody>
</table>

In this case, the missing data is uniformly distributed in the time series, and the interval between observed data is small. Thus the recovery accuracy using linear interpolation is better than the other methods. In addition, due to noniterative interpolation methods, the computation time using linear and spline interpolation is small. Due to the matrix inverse computation in each iteration, the computation time in one iteration using ALS is much larger than using ADMM. If the iterations are increasing, the computation time using ALS for the recovery will exceed the requirement of real-time applications.

### Case 2: Missing data in one channel during one second

If one channel malfunctions in one second, 25 continuous elements in one column of matrix $M$ will be missing simultaneously. A 2800-by-72 matrix with an oscillation event and a 1025-by-63 matrix in a stationary scenario are separated into $112 \times 72$ and $41 \times 63$ sub-matrices with a size of 25-by-1. We artificially assign an i.i.d. distribution of missing rate to each sub-matrices. Tables III and IV list the statistic values of convergence iterations, computational time, and RREs with different observed data probabilities in disturbance and stationary scenarios using ADMM, ALS, linear interpolation, and spline interpolation, respectively in Case 2.
it weakens the temporal correlations of the observed PMU data. Using linear interpolation, RREs are much larger than using ADMM and ALS approaches with different observed data probabilities. The spline interpolation cannot estimate the correct data in this case.

Fig. 9 shows the estimated data using ADMM, ALS, and linear interpolation methods compared to the actual values on Channel 2 from sampling instants 1676 to 1775 in the disturbed scenario, and on Channel 3 from sampling instants 50 to 125 in the stationary scenario, respectively. Since voltage magnitude measurements are oscillated drastically in both disturbed and stationary scenarios, if the data are missing during a long time, linear interpolation is difficult to estimate the correct values. On the other hand, due to the spatial correlation the recovered data are closer to the actual values using ADMM and ALS approaches.

D. Case 3: Missing data in all channels during one second

In this case, 25 successive rows of data are missing in the observed matrix $M$. The 2800-by-72 matrix contains voltage magnitude measurements with oscillation event and the 1025-by-63 matrix with measurements in stationary scenario can be treated as 112 and 41 sub-matrices with sizes of 25-by-72 and 25-by-63, respectively. The observed data probability denotes the proportion of the observed sub-matrices to the total 112 and 41 ones, respectively.

The observed matrix $M_1$ and $M_2$ are reshaped using OCCRM. The original 2800-by-72 matrix $M_1$ is reshaped to a 80-by-2520 matrix $\tilde{M}_1 (n^* = 35)$, and the original 1025-by-63 matrix $M_2$ is reshaped to a 205-by-315 matrix $\tilde{M}_2 (n^* = 5)$. The approximate ranks of the corresponding reshaped measurement matrices $\tilde{X}_1$ and $\tilde{X}_2$ are 21 and 23, respectively, with $\beta = 0.993$.

Tables V and VI list the statistic values of convergence iterations, computational time, and RREs with different observed data probabilities in disturbance and stationary scenarios using ADMM and ALS in Case 3, respectively.

Since the approximated ranks of the reshaped observed matrices are larger than the original ones, RREs using ADMM and ALS are larger than the ones in Cases 1 and 2. The rank of filled reshaped matrix with PMU data is estimated to be 20 using ALS, which results in a worse recovery accuracy compared to using ADMM. As 25 successive data samples are lost in one single channel, RRE using linear interpolation method is larger than ADMM and ALS. Using spline interpolation the missing data cannot be estimated. Compared to the computation time in one iteration using ADMM in stationary scenario in Case 2, the computation time in Case 4 is longer. The reason is that computational complexity using ADMM is $\mathcal{O}(n_1 n_2^2)$, $z = \min\{n_1, n_2\}$. In Case 2, the size of $M$ is $1025 \times 63$, and in Case 3 the size is $205 \times 315$. The computational complexity is increasing using ADMM with reshaped observed matrix.

We summarize some properties of the recovery using ADMM and ALS in Table VII.

VII. CONCLUSION

In this paper, we formulated missing PMU data recovery as a low-rank matrix completion problem, and presented an ADMM based algorithm for solving it. We illustrated our results using ADMM with real PMU measurements from Jiangsu Electric Power Company. Based on the temporal and
spatial correlations of the PMU data, ADMM approach can estimate the measurements successfully in both steady-state and disturbance scenarios. Compared to the existing ALS algorithm, the computational complexity and time are much lower. Due to the less computation time, ADMM for the recovery can be used for both off-line and online applications.

In addition, the ADMM algorithm avoids to estimate the rank of data, which reduces the influence of the uncertain factor into the performance. Also we provided algorithms of reshaping the observed matrix for recovering the lost data from all the channels during one or more sampling instants. Compared to the interpolation method, the applicable scenarios using ADMM are more various.

Our future work will include developing the optimal reshaping matrix method for improving the recovery accuracy, and recovering low-rank matrices from corrupted observations.

### References


### Table VII

<table>
<thead>
<tr>
<th></th>
<th># iterations</th>
<th>time of one iteration (s)</th>
<th>computational complexity</th>
<th>estimates at first several iterations</th>
<th>effect of the filled matrix rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALS</td>
<td>tens</td>
<td>large</td>
<td>$O(r^2(n_1 + n_2))$</td>
<td>convergence</td>
<td>sensitivity</td>
</tr>
<tr>
<td>ADMM</td>
<td>hundreds</td>
<td>small</td>
<td>$O(n_1 n_2 z)$</td>
<td>fluctuation</td>
<td>bluness</td>
</tr>
</tbody>
</table>


